Improved RJMCMC point process sampler for object detection on images by simulated annealing

Mathias Ortner, Xavier Descombes, Josiane Zerubia

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Improved RJMCMC point process sampler for object detection on images by simulated annealing

Mathias Ortner, Xavier Descombes, Josiane Zerubia

N° 4900

August 2003

THÈME 3

Rapport de recherche

ISSN 0249-6399 ISRN INRIA/RR--4900--FR+ENG

Improved RJMCMC point process sampler for object detection on images by simulated annealing

Mathias Ortner, Xavier Descombes, Josiane Zerubia

Thème 3 — Interaction homme-machine,
images, données, connaissances
Projet Ariana

Rapport de recherche n° 4900 — August 2003 — 60 pages

Abstract: We first recall Geyer and Møller algorithm that allows to sample point processes using a Markov chain. We also recall Green's framework that allows to build samplers on general state spaces by imposing reversibility of the designed Markov chain.

Since in our image processing applications, we are interested by sampling highly spatially correlated and non-invariant point processes, we adapt these ideas to improve the exploration ability of the algorithm. In particular, we keep the ability of generating points with non-uniform distributions, and design an updating scheme that allows to generate points in some neighborhood of other points.

We first design updating schemes under Green's framework to keep $\pi(.)$ reversibility of the Markov chain and then show that stability properties are not loosed. Using a drift condition we prove that the Markov chain is geometrically ergodic and Harris recurrent.

We finally show on experimental results that these kinds of updates are usefull and propose other improvements.

Key-words: MCMC point process sampler, RJMCMC, inhomogeneous Poisson point process, Geyer and Møller birth or death algorithm, birth or death in a neighborhood, drift condition

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles, BP 93, 06902 Sophia Antipolis Cedex (France)
Téléphone : +33 4 92 38 77 77 — Télécopie : +33 4 92 38 77 65

Echantillonnage de processus ponctuels par RJMCMC dans un but de détection d'objets sur une image par recuit simulé

Résumé : Nous commençons par résumer l'algorithme de Geyer et Møller qui permet, en utilisant une chaîne de Markov, d'échantillonner des lois de processus ponctuels. Nous rappelons également le cadre théorique proposé par Green qui permet d'imposer la réversibilité d'une chaîne de Markov sous une loi désirée.

Dans le cadre de nos applications en traitement d'image, nous sommes intéressés par la simulation de processus ponctuels dont la loi dépend fortement de la localisation géographique des points. Nous présentons donc ici des noyaux de proposition qui améliorent la capacité de l'algorithme de Geyer et Meyer à explorer les bons endroits de l'espace d'état. En particulier, nous proposons une transformation qui permet de faire apparaître ou disparaître des points dans un voisinage quelconque d'un autre point. Nous gardons également la possibilité de générer des points suivant une loi non uniforme.

Nous construisons donc de tels noyaux de perturbations grâce au travail de Green de manière à garder la-$\pi(.)$ réversibilité de la chaîne de Markov construite. Nous démontrons ensuite les bonnes propriétés de stabilité qui assurent le bon comportement asymptotique de la chaîne. En particulier, grâce à une condition de "drift", nous montrons l'ergodicité géométrique et la récurrence de la chaîne au sens de Harris.

Nous concluons en validant par l'expérience nos résultats théoriques, et en montrons leur utilité sur un exemple concret.

Nous proposons d'ultimes améliorations pour conclure.

Mots-clés : échantillonage MCMC de processus ponctuel, MCMC à sauts réversibles, processus ponctuels de Poisson inhomogènes, Algorithme de naissance ou mort de Geyer et Møller, naissance ou mort dans un voisinage, condition de "drift"

Contents

Introduction **4**

1 Notations **6**

2 Point process framework **7**

3 Basic Monte Carlo sampler for point process **8**
3.1 Geyer and Møller algorithm : description and properties . . . 8
3.2 Objectives of this work . . . 10
3.3 Metropolis Hastings Green . . . 10
3.3.1 Generalities . . . 10
3.3.2 State dependent mixing . . . 11
3.3.3 π-reversibility . . . 12
3.3.4 Comment . . . 13

4 Improved Proposition kernels **14**
4.1 Birth or death kernel . . . 14
4.2 Non jumping transformations . . . 16
4.2.1 General framework . . . 16
4.2.2 Simple perturbations . . . 18
4.2.3 Pre-explorative updating scheme . . . 19
4.2.4 Gibbs like sampling . . . 23
4.3 Birth or death in a Neighborhood . . . 24
4.3.1 Framework . . . 24
4.3.2 Toy example . . . 29

5 Improved MCMC point process sampler **31**
5.1 Algorithm . . . 31
5.1.1 Birth or death . . . 31
5.1.2 Non jumping transformations . . . 31
5.1.3 Birth or death in a neighborhood . . . 32
5.2 Algorithm . . . 33
5.3 Conditions . . . 33
5.3.1 Usual stability condition . . . 33
5.3.2 Mixture conditions . . . 34
5.3.3 Bounds on birth or death in a neighborhood . . . 34
5.3.4 Conditions on the relation-ship . . . 35
5.4 Property . . . 36
5.5 Proof-$\pi(.)$ reversibility . . . 36
5.6 Proof-Stability properties . . . 36
5.6.1 φ-Irreducibility and small sets . . . 36

5.6.2 Harris recurrence and geometric ergodicity . . . 38

6 Experiments **43**
6.1 Poisson point process . . . 43
6.2 Target distribution and Total Variation . . . 43
6.3 Birth or death . . . 44
6.4 Birth and death in a neighborhood . . . 44
6.4.1 Non homogeneous Poisson point process . . . 47
6.4.2 Experiments . . . 47
6.4.3 Distributions of interest . . . 48
6.4.4 Birth or death in a neighborhood . . . 48
6.4.5 Results . . . 48

7 Improved birth and death in a neighborhood **53**
7.1 Asymmetric relationships . . . 53
7.2 Pre-explorative birth and death in a neighborhood . . . 54

Conclusion **57**

Acknowledgments **58**

References **59**

Introduction

In our applications we use point processes to detect an unknown number of objects in images. This kind of approach were introduced by Van Lieshout and Baddeley in [1]. Rue in [17], [18] has also used them and explored a variety of problematics linked to image processing.

Basically, the idea underlying this type of approach is to view an image as a realization of a point process (see[5], [9] , [11] or [13] to have several applied examples in image processing). The method uses a simulated annealing and a point process sampler.

The quality of the result and the speed of the algorithm deeply depend on the sampler used. When doing statistical inference on spatial point process, it is possible to use the following algorithms, which are the main ones :

- birth and death algorithm,
- perfect simulation,
- reversible jump MCMC (discrete time),
- continuous time sampler.

See [19] to have a presentation of the first one, [10] or [20] for recent works on the second one, and [4] for the last one. In this report, we focus on the third one.

A huge literature already exists on RJMCMC and the improvements it is possible to make, in term of correlation of the designed Markov chain. When working with simulated annealing, problematic is a bit different. In [3] for instance, Brooks presents general ideas to improve RJMCMC when doing classical model selection by simulated annealing. Since our application is very specific (object detection in images) we however need other improvements.
When working with spatial statistics, point processes are quite often supposed to be spatially invariant (i.e., its distribution is invariant under translation and rotation), which is obviously not the case when doing object detection.

That is why we look at specific improvements. Here we focus on object detection, and present some ideas adapted to our framework which was presented in [13], [14], [15] and [16].

1 Notations

$S \subset \mathbb{R}^d$	bounded Borel set on which the point process is defined ($d \geq 2$)
$u \in S$	a point of S
K	subset of $\mathbb{R}^2$
M	space of marks : $S = K \times M$
$\lambda_d(.)$	Lebesgue measure on $\mathbb{R}^d$
$\mathbf{x}$	a configuration of points of S : $\mathbf{x} = \{u_1, \ldots, u_{n(\mathbf{x})}\} \quad u_i \in S$
$n(\mathbf{x})$	number of points in $\mathbf{x}$
$\mathcal{C}$	space of all possible configurations of points. (Also denoted N^{lf})
$\mathcal{N}^{lf}$	σ-algebra on $\mathcal{C}$
$\nu(.)$	measure on S : intensity measure of a Poisson point process
$\mu(.)$	distribution of a Poisson point process (measure elements from $\mathcal{N}^{lf}$).
$h(.)$	density of a point process with respect to a Poisson point process distribution.
χ	state space of some basic Markov chain
$\mathcal{B}$	σ algebra associated to χ
$(X_n)_{n\geq 0}$	Markov chain on χ
$\pi(.)$	target distribution on defined on χ
$Q(x,.)$	stochastic or sub-stochastic Kernel defined on χ, $x \in \chi$
$P(.,.)$	transition kernel of a Markov chain
$\sim$	symmetric relationship defined on S (ex: $u \sim v$)
$V(u)$	neighborhood of a point u with respect to $\sim$
$\mathcal{R}(\mathbf{x})$	set of interacting pairs of point from $\mathbf{x}$ with respect to $\sim$

2 Point process framework

Let S be a bounded and closed subset of $\mathbb{R}^d$. We assume that S can be written as the product $S = K \times M$, where $K \subseteq \mathbb{R}^2$.

We denote by u, v, w some elements of S, and $\mathbf{x} = \{u_1, \ldots, u_{n(\mathbf{x})}\}$ (resp. $\mathbf{y}$) some finite configurations of points of S.

The set of all finite configurations of points of S is denoted by $\mathcal{C}$(or N^{lf}). Details on N^{lf} and the associated σ-algebra $\mathcal{N}^{lf}$ can be found in [19].

We are interested in simulating a point process with THE distribution $\pi(.)$ on $\mathcal{C}$. This distribution is defined by a density $h(.)$ and a reference distribution $\mu(.)$. This reference distribution is the distribution of a Poisson point process with intensity measure $\nu(.)$ on S.

Figure 1a presents a realization of a uniform Poisson point process with mean 50 on $\mathbb{R}^2$, while figure 1b presents a realization of point process of rectangles with mean 20. In our

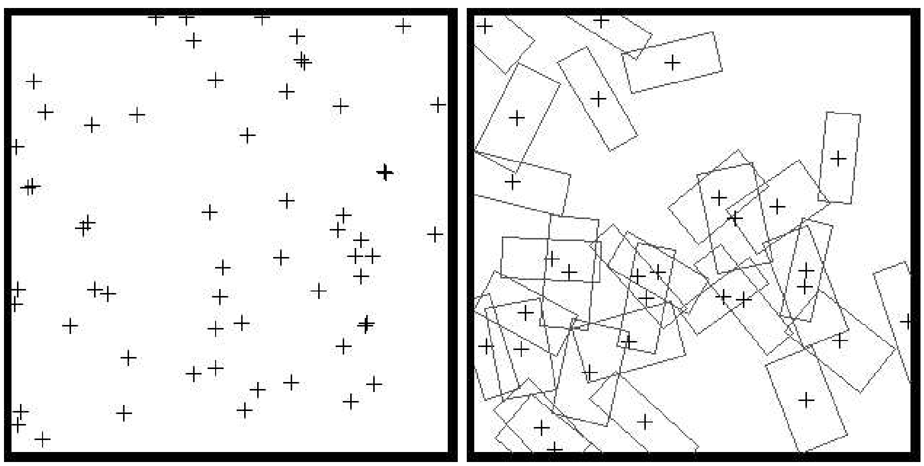


Figure 1: Two realizations, left : of a Poisson point process of points with mean 50 and right : of a Poisson point process of rectangles with mean 20.

image processing application, $h(.)$ is made of a prior information on patterns of objects and a data term that models likely images given a configuration of objects.

3 Basic Monte Carlo sampler for point process

This section relies on Geyer's work [6]. In his paper, Geyer presents a basic algorithm to sample point process distributions. In particular, he shows the relationship between the work of Green [8] and the algorithm he proposed with Møller in 1994 [7].

In this report, we extend his results in order to improve point process samplers to be able to deal with our specific applications (see [14] for instance).

We first recall here some useful concepts and notations from Geyer's work.

3.1 Geyer and Møller algorithm : description and properties

Geyer and Moller algorithm builds a Markov Chain $(X_n)_{n\geq 0}$ defined on the space $\mathcal{C}$ of finite configurations of points of S as follows :

Algorithm A

For a given state $X_t = \mathbf{x}$, with probability $\frac{1}{2}$ propose to add a point to the current configuration, and with probability $\frac{1}{2}$, propose to remove a point of the current configuration, except if $X_t = \emptyset$ in which case $X_{t+1} = X_t$:

Birth : *Generate a new point $u \in S$ according to $\frac{\nu(.)}{\nu(S)}$, propose $\mathbf{y} = \mathbf{x} \cup u$, compute*

$$R = \frac{h(\mathbf{y})}{h(\mathbf{x})}\frac{\nu(S)}{n(\mathbf{y})}$$

and accept $X_{t+1} = \mathbf{y}$ with probability $\alpha = \min(1, R)$.

Death : *Choose v uniformly in $\mathbf{x}$, propose $y = \mathbf{x} \setminus v$, compute*

$$R = \frac{h(\mathbf{y})}{h(\mathbf{x})}\frac{n(\mathbf{x})}{\nu(S)}$$

and with probability $\alpha = \min(1, R)$ accept the proposition.

This algorithm was designed by Geyer and Møller using $\nu(.)$ proportional to Lebesgue measure on S. Now, consider the following stability condition :

Condition 1 **Stability condition**

A point process with unnormalized density $h(.)$ with respect to $\mu(.)$ is stable if there exists a real number R_h such that :

$$h(\mathbf{x} \cup u) \leq R_h h(\mathbf{x}) \quad \forall \mathbf{x} \in \mathcal{C}, \quad \forall u \in S$$

This algorithm and this condition are coupled in Geyer's results in the following two propositions :

Proposition 1 *If the unnormalized density satisfies condition 1, then algorithm A :*

- *simulates a $\varphi(.)$ irreducible Markov chain and every bounded set is small,*
- *simulates a Markov chain that is* **Harris recurrent** *and* **geometrically ergodic**

This proposition gives the convergence of the Markov chain. The target distribution is given by the following proposition :

Proposition 2 *Algorithm A build a Markov chain that is $\pi(.)$ invariant, $\pi(.)$ being the distribution on $\mathcal{C}$ defined by :*

- *the reference Poisson point process distribution $\mu(.)$,*
- *the unnormalized density function $h(.)$.*

Basically, these results mean that if $(X_n)_{n\geq 0}$ is $\pi(.)$ invariant, then starting from any point of $\mathcal{C}$, it converges to $\pi(.)$ in total variation with a geometrical rate :

$$\exists r > 1 \qquad \sum_{n=1}^{\infty} r^n \|K^n(x,.) - \pi(.)\|_{TV} < \infty \qquad \forall \mathbf{x} \in \mathcal{C}$$

The quality of such a result can be seen on asymptotic theorems (see [6]). For a regular enough (Lyapounov condition) function g, empirical estimation of $\mathbb{E}_\pi(g(X))$ follows a central limit theorem, even if all points of the Markov chain are taken into account, including original states $X_0, X_1, \dots$.

3.2 Objectives of this work

The goal of this work is to add other updating schemes in the Monte Carlo Markov Chain sampler, in order to obtain some good mixing properties of $(X_n)_{n\geq 0}$ while keeping the previously described good properties of convergence.

We divide these properties of interest of the Markov Chain in two parts :

- first, **stability properties** : **irreductibility**, **harris Recurrence** and **geometric ergodicity**,

- second, $\pi(.)$ **invariance** that is obtained in the following framework by imposing $\pi(.)$ reversibility of $(X_n)_{n\geq 0}$ (the latter implying the former.)

We design some new updating schemes in section 4 insuring that the π reversibility is kept and show stability properties in section 5. Section 6 presents some experiments validating our results.

3.3 Metropolis Hastings Green

3.3.1 Generalities

Suppose we want to sample according to a distribution $\pi(.)$ known up to a normalizing constant on a space χ. One solution is to build a Markov chain that is invariant under the desired distribution. In [8], Green presents a very general framework to do this. His algorithm is a generalization of several well known ones, like for instance Metropolis-Hastings samplers.

Green proposed to build a Markov Chain $(X_n)_{n\geq 0}$ using :

- the target distribution $\pi(.)$ defined on χ and known up to a normalizing constant,
- a proposition Kernel $Q(\mathbf{x}, .)$,
- a symmetric measure $\xi(.)$ defined on $\chi \times \chi$.

If we assume that $\xi(dx, dy)$ dominates $\pi(dx)Q(x, dy)$ and note $f(x, y)$ the associated Radon-Nykodim derivative, the following updating procedure gives a Markov chain $(X_n)_{n\geq 0}$ that is $\pi(.)$ invariant :

If $X_t = x$,

1. generate $y \sim Q(x, .)$,
2. compute Green's ratio
$$R = \frac{f(y, x)}{f(x, y)}$$
3. compute the acceptance rate $\alpha = \min(1, R)$ and :
   - with probability α, accept the proposition : $X_{t+1} = y$,
   - with probability $1 - \alpha$, reject the proposition : $X_{t+1} = x$.

The $\pi(.)$ invariance of the Markov Chain is easily proved by checking that the transition kernel $P(.,.)$ of the Markov Chain is $\pi(.)$ reversible, that is to say :

$$\forall A \subseteq \chi,\ B \subseteq \chi \qquad \int_A \pi(dx) \int_B P(x, dy) = \int_B \pi(dx) \int_A P(x, dy) \tag{1}$$

3.3.2 State dependent mixing

Green actually proposed to use a substochastic mixture of proposition kernels :

$$Q(x, A) = \sum_m Q_m(x, A) \quad \text{with} \quad Q(x, \chi) \leq 1$$

Now, we assume that the following condition hold :

Hypothesis 1 *For all m, there exists a symmetric measure $\xi_m(dx, dy)$ defined on $\chi \times \chi$ dominating $\pi(dx)Q_m(x, dy)$.*

Associated Radon-Nykodim derivative is denoted by $f_m(.,.)$. For each m :

$$f_m(x, y) = \frac{\pi(dx)Q_m(x, dy)}{\xi(dx, dy)} \tag{2}$$

If we note $p_m(x) = Q_m(x, \chi)$ the probability of choosing the proposition kernel m while in state $X_t = x$, the update scheme becomes :

Starting from a given state x :

1. with probability $p_m(x)$ choose a proposition kernel Q_m. With probability $1-\sum_m p_m(x)$, set $X_{t+1} = x$.
2. simulate y according to the normalized chosen kernel :
$$y \sim \frac{Q_m(x,.)}{Q_m(x,\chi)}$$
3. Compute Green's ratio R and the acceptance rate α :
$$R_m(x,y) = \frac{f_m(y,x)}{f_m(x,y)} \qquad \alpha_m(x,y) = \min(1, R_m(x,y))$$
4. with probability α_m accept the proposition : $X_{t+1} = y$, otherwise reject it.

3.3.3 π-reversibility

We note $I(x,.)$ the identity kernel defined by $I(x,A) = \mathbf{1}_A(x)$. We suppose that the transition kernel can be written as a sum :

$$P(x,A) = d(x)I(x,A) + \sum_m P_m(x,A) \quad \text{with} \quad P(x,\chi) = 1$$

The first term $\tilde{P}(x,A) = d(x)I(x,A)$ is reversible with respect to any distribution $\pi(.)$:

$$\begin{array}{rcl} \int_A \int_B \pi(dx)\tilde{P}(x,dy) & = & \int_A \pi(dx)\ d(x)\ \mathbf{1}_B(x) \\ & = & \int \pi(dx)\ d(x)\ \mathbf{1}_A(x)\ \mathbf{1}_B(x) \\ & = & \int_B \int_A \pi(dx)\tilde{P}(x,dy) \end{array}$$

It is sufficient to show that each $P_m(.,.)$ is $\pi(.)$-reversible. In our setup, the non-identity part of the transition kernel is given by :

$$P_m(x,A) = \int_A Q_m(x,dy)\alpha_m(x,dy) \tag{3}$$

Using the hypothesis 1, we can write :

$$\int_A \int_B \pi(dx)Q_m(x,dy)\alpha_m(x,y) = \int_A \int_B f_m(x,y)\alpha_m(x,y)\xi_m(dx,dy) \tag{4}$$

Moreover, the definition of $\alpha_m(x,y)$ gives

$$f_m(x,y)\alpha_m(x,y) = f_m(y,x)\alpha_m(y,x)$$

This property and the symmetry of $\xi_m(.,.)$, give when re-injected in equation 4, the reversibility of each kernel P_m, and lead to π-reversibility of $(X_n)_{n\geq 0}$.

3.3.4 Comment

The previous proof of $\pi(.)$ reversibility underlines that "identity" parts of proposition kernels are not important, because they do not act on the invariant measure of the Markov chain. The following example show that this point is important in practice.

In the point process framework, an essential proposition kernel is the one that randomly add or delete a point from the current configuration.

This birth or death kernel consists in

1. randomly selecting birth or death (with probability 0.5 for example),
2. and then applying the chosen transformation to the current state.

If the current configuration is empty, and death has been chosen, it is allowed to propose to stay in the current state because of the identity part of the transition kernel. That is the way Geyer presented his algorithm (see [6]).

An other possibility, presented by Green (see [8]) is to adapt the probability of proposing a birth (resp. a death) to the desired discrete Poisson distribution. Of course, care is needed : when there is only one point in the current state, if death is chosen, the probability of coming back is 1, instead of 0.5 like in Geyer's framework. Taking into account exceptions in each state can become heavy when implementing the sampler.

It is of great interest to be allowed to stay in the current state when an exception is encountered, since it avoids to take into account exceptions at each step.

4 Improved Proposition kernels

We present here some proposition kernels we have implemented to improve the convergence speed in the context of our image processing application (see [14]). We focus in this section on $\pi(.)$-reversibility.

Using the previously exposed framework, we compute acceptance ratio that insure $\pi(.)$ reversibility of the designed Markov chain.

We first propose an alternative proof of $\pi(.)$ reversibility of the birth and death kernel, using L. Garcin's work [5]. Then we present some particular proposition kernels, like pre-explorative scheme to approximate marginal distributions or birth and death in a neighborhood.

4.1 Birth or death kernel

We have seen how to obtain $\pi(.)$ reversibility for a Markov chain. To describe birth or death transformation, both Geyer and Green use an infinite mixture of proposition kernels, Q_m, each acting only on configurations of m points, with a probability of choosing Q_m being equal to 1 if $n(\mathbf{x}) = m$ and 0 otherwise. Then reversibility of each of these kernels was shown.

Here we present an other proof (we initially presented in [13] and [14]) that uses only one kernel acting on any configuration of points.

Proof of $\pi(.)$ reversibility

We suppose that birth generates a point in S according to the probability distribution $\frac{\nu(.)}{\nu(S)}$ and that death uniformly chooses a point in the current configuration. We first write this kernel as :

$$q(x,.) = p_b(x)Q_b(x,.) + p_d(x)Q_d(x,.)$$

with the two kernels of birth and death defined by :

$$Q_b(\mathbf{x}, A) = \int_{u\in S} \mathbf{1}_A(\mathbf{x}\cup u)\frac{\nu(du)}{\nu(S)} \qquad A \in \mathcal{N}_s^{lf}$$

and

$$Q_d(\mathbf{x}, A) = \sum_{u\in\mathbf{x}} \mathbf{1}_A(\mathbf{x}\setminus u)\frac{1}{n(\mathbf{x})}$$

except if $n(x) = 0$ in which case $Q_d(x,.) = I(x,.)$. In the following, this identity part is forgotten. We then consider the measure ξ, where A and B are $\mu(.)$ measurable subsets of

$\mathcal{C}$, as :

$$\xi(A \times B) = \int_{\mathcal{C}} \int_{u \in S} \mathbf{1}_A(\mathbf{x})\mathbf{1}_B(\mathbf{x} \cup u)\nu(du)\mu(d\mathbf{x}) + \int_{\mathcal{C}} \mathbf{1}_A(\mathbf{x}) \sum_{u \in \mathbf{x}} \mathbf{1}_B(\mathbf{x} \setminus u)\mu(d\mathbf{x}) \quad (5)$$

We need to show that this measure is symmetric. This comes from the fact that $\nu(.)$ is the intensity measure of the Poisson process law $\mu(.)$. Let us take $A_n = A \cap N_n^f$ where N_n^f is the subset of N_S^{lf} made of configurations of n points. The definition of a Poisson process gives (see [19] for details) :

$$\begin{aligned}
\xi(A_n \times B_{n-1}) &= \frac{e^{-\nu(S)}}{n!} \int_{S^n} \sum_{u \in \mathbf{x}} \mathbf{1}_{A_n}(\mathbf{x})\mathbf{1}_{B_{n-1}}(\mathbf{x} \setminus u)\nu^n(d\mathbf{x}) && \text{(6a)}\\
&= \frac{e^{-\nu(S)}}{n!} \int_{S^n} n\mathbf{1}_{A_n}(\{x_1, \ldots, x_n\})\mathbf{1}_{B_{n-1}}(\{x_1, \ldots, x_n\})\nu^n(d\mathbf{x}) && \text{(6b)}\\
&= \frac{e^{-\nu(S)}}{(n-1)!} \int_{S^{n-1}} \int_S \mathbf{1}_{B_{n-1}}(\mathbf{y})\mathbf{1}_{A_n}(\mathbf{y} \cup u)\nu^{n-1}(d\mathbf{y})\nu(du) && \text{(6c)}\\
&= \xi(B_{n-1} \times A_n) && \text{(6d)}
\end{aligned}$$

It is useful to observe in equations (6) that if A is made of configurations that cannot be obtained by adding or removing a point from any configuration of B, the measure $\xi(A, B)$ is null.

Using equation (6), symmetry of ξ is obtained by computing $\xi(A, B)$ as an infinite sum $\sum \xi(A_n, B_{n-1}) + \xi(A_n, B_{n+1})$.

Now, we need to show that $\xi(d\mathbf{x}, d\mathbf{x}')$ dominates $\pi(d\mathbf{x})Q(\mathbf{x}, d\mathbf{x}')$ and to calculate the Radon-Nikodym derivative.

It is obvious that if a set $A \times B$ has a strictly positive $\pi(.)Q(.;.)$ measure, its ξ measure is also strictly positive.

Thus, we have two cases to consider :

1. if $\mathbf{y} = \mathbf{x} \cup u$, then expressions of π and Q give :

   $$\pi(d\mathbf{x})Q(\mathbf{x}, d\mathbf{y}) = h(\mathbf{x})\mu(d\mathbf{x})p_b(\mathbf{x})\frac{\nu(du)}{\nu(S)}$$

   And the definition of ξ (cf. equation (5)) gives :

   $$\xi(d\mathbf{x}, d\mathbf{y}) = \mu(d\mathbf{x})\nu(du)$$

   from which follows absolute continuity and Radon-Nikodym derivative :

   $$\boxed{f(\mathbf{x}, \mathbf{y}) = p_b(\mathbf{x})\frac{h(\mathbf{x})}{\nu(S)}}$$

2. The other case consists in considering : $\mathbf{y} = \mathbf{x} \setminus u$.

$$\pi(d\mathbf{x})Q(\mathbf{x}, d\mathbf{y}) = h(\mathbf{x})\mu(d\mathbf{x})p_d(\mathbf{x})\frac{1}{n(\mathbf{x})}$$

   The measure ξ gives :

$$\xi(d\mathbf{x}, d\mathbf{y}) = \mu(d\mathbf{x})$$

   and the derivative :

$$\boxed{f(\mathbf{x}, \mathbf{y}) = p_d(\mathbf{x})\frac{h(\mathbf{x})}{n(\mathbf{x})}}$$

We may infer Green's ratio from equations (1) and (2). This coefficient has two different expressions, depending on how $\mathbf{y}$ is obtained (ie. by adding or removing a point in $\mathbf{x}$) :

1. In case of a **birth**, $\mathbf{y} = \mathbf{x} \cup u$, and Green's Ration is given by :

$$\boxed{R(\mathbf{x}, \mathbf{y}) = \frac{f(\mathbf{y}, \mathbf{x})}{f(\mathbf{x}, \mathbf{y})} = \frac{p_d(\mathbf{y})}{p_b(\mathbf{x})}\frac{h(\mathbf{y})}{h(\mathbf{x})}\frac{\nu(S)}{n(\mathbf{y})}} \tag{7}$$

   where $n(\mathbf{y})$ can be replaced by $n(\mathbf{x}) + 1$.

2. In case of a **death**, $\mathbf{y} = \mathbf{x} \setminus u$, and Green's Ratio is given by :

$$\boxed{R(\mathbf{x}, \mathbf{y}) = \frac{f(\mathbf{y}, \mathbf{x})}{f(\mathbf{x}, \mathbf{y})} = \frac{p_b(\mathbf{y})}{p_d(\mathbf{x})}\frac{h(\mathbf{y})}{h(\mathbf{x})}\frac{n(\mathbf{x})}{\nu(S)}} \tag{8}$$

This gives Green's ratio for the birth or death sub-kernel.

4.2 Non jumping transformations

4.2.1 General framework

These transformations do not add or remove one point : they only change some parameters of a randomly chosen point. Transformations like translation or rotation are elements belonging to this class of transformation. This kind of transformation can be seen as a usual Metropolis Hastings updating scheme. However, we present here a framework that allows :

- to perturb a point chosen according to a state dependent distribution,
- and to use some pre-exploration step to do Gibbs-like updating scheme.

Notations

We consider a set Σ, and an associated measure $s(.)$. Usually, Σ will be a subset of $\mathbb{R}^k$ or a countable space and s the associated Lebesgue measure or the countable one.

For a given configuration of points $\mathbf{x}$ and a given point u in $\mathbf{x}$, we consider a random variable $Z_{(\mathbf{x},u)}$ living on a set $\Sigma(\mathbf{x},u) \subset \Sigma$.

We note $\mathbb{P}_Z^{(\mathbf{x},u)}(.)$ the distribution of $Z_{(\mathbf{x},u)}$ on $\Sigma(\mathbf{x},u)$, and we assume that it is dominated by $s(.)$. We denote by $f_Z^{(\mathbf{x},u)}(.)$ the associated density function.

We then consider an **injection** $\zeta_{\mathbf{x}}$:

$$\begin{array}{rccc} \zeta_{\mathbf{x}}: & S \times \Sigma & \rightarrow & S \\ & (u,z) & \rightarrow & v \end{array} \tag{9}$$

Proposition Kernel

The sub-kernel used can be described as follows :

The current state being $\mathbf{x} = \{u_1, \ldots, u_{n(\mathbf{x})}\}$,

1. choose u among the u_i with a discrete probability law $j^{\mathbf{x}}(u_i)$,
2. generate z using the distribution of $Z_{(\mathbf{x},u)}$,
3. compute v as : $v = \zeta_{\mathbf{x}}(u, Z)$,
4. propose $\mathbf{y} = \mathbf{x} \setminus u \cup v$.

This proposition kernel can be written as :

$$Q(\mathbf{x}, A) = \sum_{u \in \mathbf{x}} j^{\mathbf{x}}(x,u) \mathbb{P}_Z^{(\mathbf{x},u)}(\mathbf{1}_A(\mathbf{x} \setminus u \cup \zeta_{\mathbf{x}}(u,z)) \tag{10}$$

We consider the following measure ξ :

$$\xi(A \times B) = \int_A \sum_{u \in \mathbf{x}} \int_{\Sigma(\mathbf{x},u)} \mathbf{1}_B(\mathbf{x} \setminus u \cup \zeta_x(u,z)) s(dz) \mu(d\mathbf{x}) \tag{11}$$

To achieve symmetry of ξ we assume the symmetry of the transformation

$$v = \zeta_{\mathbf{x}}(u,z) \iff \exists \tilde{z} \in \Sigma(\mathbf{y},v) \text{ s.t. } \mathbf{y} = \mathbf{x} \setminus u \cup v \quad u = \zeta_{\mathbf{y}}(v, \tilde{z}) \tag{12}$$

and the $s(.)$ equality of the corresponding spaces

$$s(dz) = s(d\tilde{z}) \tag{13}$$

since unicity of $\tilde{z}$ comes from injectivity of $\zeta_x(u,.)$.

We thus obtain the following Green's Ratio :

$$\boxed{R(\mathbf{x},\mathbf{y}) = \frac{h(\mathbf{y})}{h(\mathbf{x})}\frac{j^{\mathbf{y}}(v)}{j^{\mathbf{x}}(u)}\frac{f_Z^{(\mathbf{y},v)}(\tilde{z})}{f_Z^{(\mathbf{x},u)}(z)}} \tag{14}$$

This recalls the ratio of usual Hastings Metropolis update.

The condition given by equation (13) is obtained easily if :

- $\Sigma \subseteq \mathbb{R}^k$ and $s(.)$ is a Lebesgue measure,
- Σ is a countable space and $s(.)$ the corresponding countable measure,
- Σ is a product of such sets and $s(.)$ product of the associated measures.

4.2.2 Simple perturbations

We present here simple transformations in the case of configuration of rectangles : $M \subset \mathbb{R}^3$ (angle, length and width).

Translation

For this transformation $\zeta_{(\mathbf{x},u)}$ does not depend on $\mathbf{x}$ and $\Sigma = [-\delta x, \delta x] \times [-\delta y, \delta y]$. We then use

$$\zeta(.;z) : u = \begin{pmatrix} x \\ y \\ l \\ L \\ \theta \end{pmatrix} \rightarrow v = \begin{pmatrix} x + z_x \\ y + z_y \\ l \\ L \\ \theta \end{pmatrix}$$

Since Σ is symmetric, condition (12) is achieved by taking $\tilde{z} = -z$.

If $(x + z_x, y + z_y) \notin K$, the proposition is rejected without any incidence on the invariant distribution, as pointed out in section 3.

In practice, we have taken uniform distributions :

$$j^{\mathbf{x}}(.) = \frac{1}{n(\mathbf{x})}\sum_{u\in\mathbf{x}} \mathbf{1}_u(.) \qquad f_Z^{\mathbf{x}}(.) = \frac{\lambda_\Sigma(.)}{\lambda_\Sigma} \qquad R(x,y) = \frac{h(\mathbf{y})}{h(\mathbf{x})}$$

Rotation

For this transformation, we use : $z \in \Sigma = [-\delta_\theta; +\delta_\theta] \subseteq \mathbb{R}$ and ζ :

$$\zeta(.;z) : u = \begin{pmatrix} x \\ y \\ l \\ L \\ \theta \end{pmatrix} \rightarrow v = \begin{pmatrix} x \\ y \\ l \\ L \\ \theta + \delta_\theta \end{pmatrix}$$

Dilations

We use two dilations, The first one is related to the length, the second to the width of a rectangle. We only explain here the first one.

We take $\Sigma = \{0, 1\} \times [-\delta_L; +\delta_L]$. Writing $z = (z_k, z_L)$ we define ζ as :

$$\zeta(.;z) : u = \begin{pmatrix} x \\ y \\ l \\ L \\ \theta \end{pmatrix} \rightarrow v = \begin{pmatrix} x + z_L \cos(\theta + z_k * \pi) \\ y + z_L \sin(\theta + z_k * \pi) \\ l \\ L + z_L \\ \theta \end{pmatrix}$$

4.2.3 Pre-explorative updating scheme

Using the framework of section 4.2.1, we present here a Gibbs-like update scheme.

The basic idea is to use a mixture of distribution on the set $\Sigma(\mathbf{x}, u)$ which is partitioned into several subsets.

Example : Rotation

Figure 2 presents the rotation transformation on a rectangle and shows the idea underlying pre-explorative updating scheme.

Using a integer N, a resolution parameter $\delta\theta$, and the angle $\theta(u)$ of the selected object u, we write :

$$\forall i \in -N, \ldots, N \quad \theta_i^u = \theta(u) + i * 2 * \delta\theta \quad \Delta\theta = (N+1) * \delta\theta$$

and

$$\Sigma(\mathbf{x}, u) = [\theta(u) - \Delta\theta, \theta(u) + \Delta\theta] = \bigcup_{i=-N}^{N} [\theta_i^u - \delta\theta, \theta_i^u + \delta\theta] = \bigcup_{i=-N}^{N} B_i^u$$

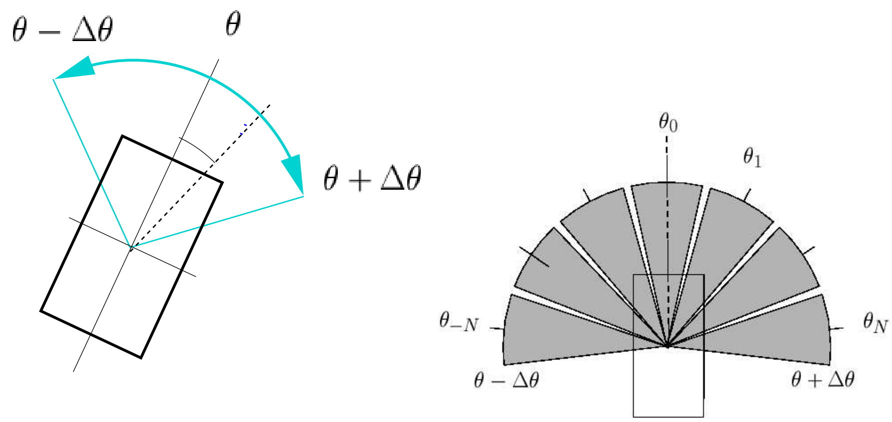


Figure 2: Rotation and pre-explorative rotation scheme.

This is a partition of $\Sigma(\mathbf{x}, u)$ into balls $B_i^u$of equal radius $\delta\theta$ centered in θ_i. For the ζ function, we use :

$$\zeta(.; z) : u = \begin{pmatrix} x \\ y \\ l \\ L \\ \theta \end{pmatrix} \rightarrow v = \begin{pmatrix} x \\ y \\ l \\ L \\ z \end{pmatrix}$$

Distribution of Z is taken as a mixture of uniform distributions on B_i^u :

$$f_Z^{(\mathbf{x},u)}(.) = \sum_{i=-N}^{N} p_i^{(\mathbf{x},u)} \frac{\mathbf{1}_{B_i}(.)}{2 * \delta\theta}$$

Now, let introduce $h(.)$ the density we want to simulate, $U(.)$ its associated energy, $\bar{h}(.)$ the **normalized** version of $h(.)$, $\tilde{T}$ a temperature parameter :

$$\bar{h}(.) = \frac{h(.)}{\int h} \qquad \bar{h}(.) \propto \exp(-U(.))$$

We write :

$$\mathbf{x}_i = \mathbf{x} \setminus u \cup \zeta(u, \theta_i) \qquad h_i(\mathbf{x}) = \bar{h}(\mathbf{x}_i)$$

We propose to take

$$p_i^{(\mathbf{x},u)} \propto h_i^{\frac{1}{\tilde{T}}}(\mathbf{x}) \qquad p_i^{(\mathbf{x},u)} = \frac{h_i^{\frac{1}{\tilde{T}}}(\mathbf{x})}{\sum_j h_j^{\frac{1}{\tilde{T}}}(\mathbf{x})} \tag{15}$$

For a suitable couple of objects (u,v), we note $I(u,v)$ the unique integer such that :

$$i = I(u,v) \iff \theta(v) \in B_i^u \tag{16}$$

The updating scheme is now :

1. choose $u \in \mathbf{x}$ randomly according to the discrete distribution $j^x(.)$.
2. compute $(\theta_{-N}, \ldots, \theta_N)$ depending on u,
3. choose one of the B_i according to the discrete distribution $(p_i^{(\mathbf{x},u)})$,
4. generate $z \in B_i$ uniformly,
5. compute $v = \zeta(u,z)$ and $\mathbf{y} = \mathbf{x} \setminus u \cup v$,
6. compute $(\theta'_{-N}, \ldots, \theta'_N)$ depending on v,
7. calculate $(p_i^{(\mathbf{y},v)})$ and especially $p_I(v,u)$, ie. the weight associated to the ball that allows to come back from v to u.
8. compute Green's ratio :
$$R(\mathbf{x},\mathbf{y}) = \frac{h(\mathbf{y})}{h(\mathbf{x})} \frac{j^{\mathbf{y}}(v)}{j^{\mathbf{x}}(u)} \frac{p_{I(v,u)}}{p_{I(u,v)}} \tag{17}$$
9. accept the update with probability $\min(1,R)$.

Now, we go further into details concerning :

- the computation of the p_i,
- the expression of R,
- and the relationship with Gibbs update.

Computation of the p_i

A first problem is related to computation of the (p_i). Their expression is given by equation 15, that is hardly usable for practical reasons. Since the (p_i) are normalized :

$$\begin{aligned} p_i^{(\mathbf{x},u)} &\propto \bar{h}^{\frac{1}{\tilde{T}}}(\mathbf{x}_i) && \text{(18a)} \\ &\propto \left(\frac{\bar{h}(\mathbf{x}_i)}{\bar{h}(\mathbf{x})}\right)^{\frac{1}{\tilde{T}}} && \text{(18b)} \\ &\propto \exp\left(-\frac{U(\mathbf{x}_i)-U(\mathbf{x})}{\tilde{T}}\right) && \text{(18c)} \end{aligned}$$

And we obtain the expression of p_i :

$$\boxed{p_i^{(\mathbf{x},u)} = \frac{\exp\left(-\frac{U(\mathbf{x}_i)-U(\mathbf{x})}{\tilde{T}}\right)}{\sum_{j=-N}^{N}\exp\left(-\frac{U(\mathbf{x}_j)-U(\mathbf{x})}{\tilde{T}}\right)}} \tag{19}$$

This expression is useful because we only need to be able to compute $2N+1$ differences of energy between two configurations $\mathbf{x}_i$ and $\mathbf{x}$ which differ only in one point.

Property

The same kind of argument applies to $\mathbf{y}$ and associated $\mathbf{y}_i$ and makes the associated Green's ratio easy to compute :

$$\begin{aligned} \tau(\mathbf{x},\mathbf{y}) &= \frac{p_{I(v,u)}}{p_{I(u,v)}} \\ &= \frac{\exp\left(-\frac{U(\mathbf{y}_{I(v,u)})-U(\mathbf{y})}{\tilde{T}}\right)}{\exp\left(-\frac{U(\mathbf{x}_{I(u,v)})-U(\mathbf{x})}{\tilde{T}}\right)}\frac{\sum_{j=-N}^{N}\exp\left(-\frac{U(\mathbf{x}_j)-U(\mathbf{x})}{\tilde{T}}\right)}{\sum_{j=-N}^{N}\exp\left(-\frac{U(\mathbf{y}_j)-U(\mathbf{y})}{\tilde{T}}\right)} \\ &= \exp\left(-\frac{U(\mathbf{y}_{I(v,u)})-U(\mathbf{x}_{I(u,v)})}{\tilde{T}}\right)\frac{\sum_{j=-N}^{N}\exp\left(-\frac{U(\mathbf{x}_j)}{\tilde{T}}\right)}{\sum_{j=-N}^{N}\exp\left(-\frac{U(\mathbf{y}_j)}{\tilde{T}}\right)} \end{aligned}$$

To give an intuitive explanation, we now assume that the ratio of the sums is equal to one. This is achieved, if for instance $(\theta^u_{-N},\ldots,\theta^u_N) = (\theta^v_{-N},\ldots,\theta^v_N)$. It is possible to obtain such an equality by imposing :

$$\Delta\theta = \pi \quad \Sigma(\mathbf{x},u) = \Sigma(\mathbf{y},v) = \cup_{i=-N}^{i=N} B(\hat{\theta}_i, \delta\theta) \tag{20}$$

using some previously fixed $\hat{\theta}_i$ as center of the B_i. This means Σ is simply subdivided into bands that do not depend neither on u nor on v.

Another simplification is obtained by putting $\tilde{T} = 1$ and using uniform distributions as $j^{\mathbf{x}}(.)$ and $j^{\mathbf{y}}(.)$. This gives :

$$R(\mathbf{x}, \mathbf{y}) = \exp\left[-\left(U(\mathbf{y}) - U(\mathbf{x}_{I(u,v)})\right) + \left(U(\mathbf{x}) - U(\mathbf{y}_{I(v,u)})\right)\right]$$

Since $\mathbf{x}_{I(u,v)}$ corresponds to the center of the ball that gives $\mathbf{y}$ by a uniform perturbation of radius $\delta\theta$, and $\mathbf{y}_{I(v,u)}$ corresponds to $\mathbf{x}$, if we make $\delta\theta$ tends to zero, R tends to one :

$$\boxed{\lim_{\delta\theta \to 0} R = 1} \tag{21}$$

It is thus possible to limit the correlation of the Markov chain by making $\delta\theta$ small enough.

4.2.4 Gibbs like sampling

It is worth seeing that under condition (20), the previously defined transformation can be considered as a discrete estimation of the marginal distribution of $\pi(.)$ with respect to the angle parameter.

From this point of view, this update scheme can be seen as a kind of Gibbs updating scheme, where the acceptance ratio corrects the discretisation error.

4.3 Birth or death in a Neighborhood

In this section we show how to make birth of a point in some defined neighborhood of another point. In our urban processing application, we use this kind of updating to make buildings appear in the alignment of other buildings, since our probability density favorites such alignments.

4.3.1 Framework

Definitions

Let $\sim$ be a symmetric relation on S. For instance :

$$u \sim v \iff \begin{cases} d_K(u,v) \leq d_{\max} \\ u \neq v \end{cases} \tag{22}$$

where $d_K(.,.)$ is the Euclidean distance on K. For a given configuration $\mathbf{x}$ define $\mathcal{R}(\mathbf{x})$ as the set of pairs of related points :

$$\mathcal{R}(\mathbf{x}) = \{\{u,v\},\ \ u \in \mathbf{x},\ \ v \in \mathbf{x}\ \text{ s.t. } u \sim v\} \tag{23}$$

$V(u) \subseteq S$ the neighborhood of a point :

$$V(u) = \{v \in S \ \text{ s.t. } \ v \sim u\} \tag{24}$$

and $V(\mathbf{x}) \subseteq S$ the neighborhood of a configuration :

$$V(\mathbf{x}) = \{u \in S \ \text{ s.t. } \ \exists v \in \mathbf{x} \ \ u \sim v\} \tag{25}$$

Kernel

Consider the following proposition **birth or death** kernel :

Birth

1. Choose u among $\mathbf{x} = \{u_1, \ldots, u_n\}$, using a probability distribution $j_b^{\mathbf{x}}(.)$,
2. generate v such that $v \in V(\mathbf{x})$,
3. propose $\mathbf{y} = x \cup v$.

Death

1. Randomly choose a point $u \in \mathbf{x}$ such that $V(u) \neq \emptyset$,

2. propose $\mathbf{y} = \mathbf{x} \setminus u$.

Such a kernel can be written as :

$$Q(\mathbf{x}, .) = p_b(\mathbf{x}) Q_b(\mathbf{x}, .) + p_d(\mathbf{x}) Q_d(\mathbf{x}, .)$$

with

$$Q_b(\mathbf{x}, A) = \sum_{u \in \mathbf{x}} j_b^{\mathbf{x}}(u) q_b^u(\mathbf{x}, A) \qquad Q_d(\mathbf{x}, A) = \sum_{u \in \mathbf{x}} j_d^{\mathbf{x}}(u) \mathbf{1}_A(\mathbf{x} \setminus u) \tag{26}$$

where obviously $j_d^{\mathbf{x}}(u)$ should be null for any u not belonging to any pair of $\mathcal{R}(\mathbf{x})$, and $q_b^u(\mathbf{x}, .)$ add points in the neighborhood of u.

Measure ξ

To detail the conditions imposed on the previous kernel, we need a symmetric measure dominating $\pi(d\mathbf{x})Q(d\mathbf{x}, d\mathbf{y})$. It is convenient to take the symmetric measure already used for usual birth and death :

$$\xi(A \times B) = \int_{\mathcal{C}} \int_{u \in S} \mathbf{1}_A(\mathbf{x}) \mathbf{1}_B(\mathbf{x} \cup u) \nu(du) \mu(d\mathbf{x}) + \int_{\mathcal{C}} \mathbf{1}_A(\mathbf{x}) \sum_{u \in \mathbf{x}} \mathbf{1}_B(\mathbf{x} \setminus u) \mu(d\mathbf{x})$$

Since birth or death in a neighborhood is a particular case of birth or death, ξ dominates $Q(., .)\pi(.)$. We now calculate the associated derivative, by considering two cases.

Birth kernel and associated derivative

First, let detail the way of generating a neighbor of an already chosen object u :

1. generate a vector z on a space Σ, according to the law of a random variable Z,
2. apply an injection $\eta_u(.)$ on z :

$$\begin{array}{rccc} & \Sigma & \rightarrow & S \\ \eta_u : & z & \rightarrow & v \end{array}$$

By this way, the couple (z, η_u) gives an object v that should be a neighbor of u thus the first condition we impose on η is that

$$\eta_u(\Sigma) = V(u) \tag{27}$$

Assuming that Z has a distribution on Σ given by $\mathbf{P}_Z$, it is possible to detail q_b :

$$\begin{aligned} q_b^u(\mathbf{x}, A) &= \mathbf{P}_Z(\mathbf{x} \cup \eta_u(Z) \in A) \\ &= \mathbf{P}_Z(\eta_u(Z) \in A_{\mathbf{x}}) \end{aligned}$$

where $A_{\mathbf{x}} \subseteq S$ corresponds to the following set :

$$A_{\mathbf{x}} = \{v \in S \quad \text{s.t.} \qquad \mathbf{x} \cup v \in A\}$$

This allows us to write :

$$q_b^u(\mathbf{x}, A) = \int_\Sigma \mathbf{1}_{A_{\mathbf{x}}}(\eta_u(z))\, d\mathbf{P}_Z(z)$$

To go further two assumptions are needed :

- $\eta(.)$ is a diffeomorphism, implying that Σ is required to have the same dimension than S (condition of **dimension matching** in Green's framework),
- the distribution of Z is dominated by Lebesgue measure and has a Radon Nikodym derivative $f_Z(.)$. It can be, for instance, the uniform distribution :

$$f_Z(.) = \frac{1}{\lambda_S(\Sigma)}$$

These assumptions allow the following 'change of variable' involving the Jacobian of $\eta_u(.)$:

$$\begin{aligned} q_b^u(\mathbf{x}, A) &= \int_\Sigma \mathbf{1}_{A_{\mathbf{x}}}(\eta_u(z))\; f_Z(z)\; \lambda(dz) && (28) \\ &= \int_{\eta_u(\Sigma)} \mathbf{1}_{A_{\mathbf{x}}}(v)\; f_Z(\eta_u^{-1}(v))\; d\lambda \bullet \eta_u(v) && (29) \\ &= \int_{\eta_u(\Sigma)} \mathbf{1}_{A_{\mathbf{x}}}(v)\; f_Z(\eta_u^{-1}(v))\; |J_{\eta_u^{-1}}(v)|\; \lambda(dv) && (30) \\ &= \int_{V(u)} \mathbf{1}_{A_{\mathbf{x}}}(v)\; f_Z(\eta_u^{-1}(v))\; |J_{\eta_u^{-1}}(v)|\; \lambda(dv) && (31) \end{aligned}$$

Let introduce Λ_u defined as follows :

$$\boxed{\Lambda_u(v) = \begin{cases} 0 & \text{if} \quad v \notin V(u) = \eta_u(\Sigma) \\ |J_{\eta_u^{-1}}(v)| & \text{otherwise} \end{cases}} \qquad (32)$$

Thus equation (31) becomes :

$$q_b^u(\mathbf{x}, A) = \int_S \mathbf{1}_{A_{\mathbf{x}}}(v)\; f_Z(\eta_u^{-1}(v))\; \Lambda_u(v)\lambda(dv) \qquad (33)$$

And

$$Q_b(\mathbf{x}, A) = \int_S \sum_{u \in \mathbf{x}} j_b^{\mathbf{x}}(u)\ \mathbf{1}_A(\mathbf{x} \cup v)\ f_Z(\eta_u^{-1}(v))\ \Lambda_u(v)\lambda(dv)$$

For two given A and B, subset of $\mathcal{C}$, define $A_n = A \cap N_n^f$ and $B_{n+1} = B \cap N_{n+1}^f$, for all n.

$$\int_{A_n} \int_{B_{n+1}} \pi(d\mathbf{x})\ Q(\mathbf{x}, d\mathbf{y}) = \int_{A_n} \int_S h(\mathbf{x})\ \mathbf{1}_{B_{n+1}}(\mathbf{x} \cup v)\ p_b \sum_{u \in \mathbf{x}} j_b^{\mathbf{x}}(u)\ f_Z(\eta_u^{-1}(v))\ \Lambda_u(v)\ \lambda(dv)\mu(d\mathbf{x}) \tag{34}$$

To identify the measure ξ in this expression, we assume that $\nu(.)$, intensity measure of the reference Poisson Point process, is dominated by the Lebesgue measure $\lambda_S(.)$, and we denote by f_ν the associated Radon Nikodym derivative :

$$\nu(A) = \int_A f_\nu(u)\lambda_S(du)$$

Equation (34) then becomes :

$$\int_{A_n} \int_{B_{n+1}} \pi(d\mathbf{x})\ Q(\mathbf{x}, d\mathbf{y}) = \int_{\mathcal{C}} \int_S \mathbf{1}_{A_n}(\mathbf{x})\ \mathbf{1}_{B_{n+1}}(\mathbf{x} \cup v)\ h(\mathbf{x})\ p_b \sum_{u \in \mathbf{x}} j_b^{\mathbf{x}}(u)\ f_Z(\eta_u^{-1}(v))\ \Lambda_u(v)\ \nu(dv)\mu(d\mathbf{x}) \tag{35}$$

The last equation allows to conclude on the Radon Nikodym derivative of interest in the case of a birth :

$$\boxed{f(\mathbf{x}, \mathbf{x} \cup v) = \frac{p_b}{f_\nu(v)}\ h(\mathbf{x})\ \left(\sum_{u \in \mathbf{x}} j_b^{\mathbf{x}}(u)\ f_Z(\eta_u^{-1}(v))\ \Lambda_u(v) \right)}$$

Death kernel and associated derivative

The death kernel expression is very simple

$$Q_d(\mathbf{x}, A) = \sum_{u \in \mathbf{x}} j_d^{\mathbf{x}}(u)\mathbf{1}_A(\mathbf{x} \setminus u)$$

with an implicit constraint on the $(j_d^{\mathbf{x}})_{u \in \mathbf{x}}$:

$$\text{If }\ u \notin V(\mathbf{x})\ \ j_d^{\mathbf{x}}(u) = 0 \tag{36}$$

Now we can compute the Radon Nikodym derivative :

$$\int_{A_{n+1}} \int_{B_n} \pi(d\mathbf{x})\ Q(\mathbf{x}, d\mathbf{y}) = \int_{\mathcal{C}} \mathbf{1}_{A_{n+1}}(\mathbf{x})\ p_d\ h(\mathbf{x}) \sum_{u \in \mathbf{x}} j_d^{\mathbf{x}}(u)\ \mathbf{1}_{B_n}(\mathbf{x} \setminus u)\ \mu(d\mathbf{x}) \tag{37}$$

And by identifying the measure ξ, it follows :

$$\boxed{f(\mathbf{x}, \mathbf{x} \setminus u) = h(\mathbf{x})\ p_d\ j_d^{\mathbf{x}}(u)}$$

Green's ratio

Using Radon Nikodym derivatives, it is possible to calculate Green's ratio :

- in the case of a birth :

$$\boxed{R(\mathbf{x}, \mathbf{x} \cup v) = \frac{h(\mathbf{x} \cup v)}{h(\mathbf{x})} \frac{p_d}{p_b} \frac{j_d^{\mathbf{x} \cup v}(v)\ f_\nu(v)}{\sum_{u \in \mathbf{x}} j_b^{\mathbf{x}}(u)\ f_Z(\eta_u^{-1}(v)) \Lambda_u(v)}} \tag{38}$$

- in the case of a death :

$$\boxed{R(\mathbf{x}, \mathbf{x} \setminus v) = \frac{h(\mathbf{x} \setminus v)}{h(\mathbf{x})} \frac{p_b}{p_d} \frac{\sum_{u \in \mathbf{x} \setminus v} j_b^{\mathbf{x} \setminus v}(u)\ f_Z(\eta_u^{-1}(v))\ \Lambda_u(v)}{j_d^{\mathbf{x}}(v)\ f_\nu(v)}} \tag{39}$$

Comments

The expressions obtained are very intuitive, since the ratio can be formulated as the likelihood ratio multiplied by the probability of proposing the coming back move divided the probability of proposing the forward perturbation. Thus, it is possible to see the sum of Jacobian as an expected Jacobian.

From a practical point of view, the computation of Green's ratio is decomposed as follows :

In case of a birth :

1. compute the discrete distribution $j_b^x(.)$ and choose u according to it,
2. generate z on Σ, compute $v = \eta_u(z)$ and $f_\nu(v)$,
3. compute the pairs from $\mathcal{R}(\mathbf{x})$ containing v, and for each pair $\{v, w\}$ compute $j_b(w)\Lambda_w(v)$ and $f_Z(\eta_u^{-1}(v))$,
4. compute the probability $j_d^{\mathbf{x} \cup v}(v)$ of choosing v from $\mathbf{x} \cup v$
5. compute $h(\mathbf{x} \cup v)/h(\mathbf{x})$,
6. compute R using its expression.

In case of a death, the procedure is almost the same :

1. choose u to be killed,
2. compute pairs and Jacobian involving u and the probabilities of choosing v related to u according to $j_b^{\mathbf{x}\setminus u}$
3. compute $h(\mathbf{x} \setminus u)/h(\mathbf{x})$.

In practice

We took

- a uniform probability distribution on the space of mark M :

$$\nu(.) = \left(\lambda_K \times \frac{\lambda_M}{\lambda_M(M)}\right) \qquad f_\nu = \frac{1}{\lambda_M(M)} \tag{40}$$

- a uniform discrete distribution over objects for birth :

$$j_b^{\mathbf{x}}(u) = \frac{1}{n(\mathbf{x})} \tag{41}$$

- a uniform discrete distribution over pairs of related objects for death :

$$j_d^{\mathbf{x}}(u) = \frac{\frac{1}{2} * \mathrm{card}\left\{\{v,w\} \in \mathcal{R}(\mathbf{x}) \qquad u \in \{v,w\}\right\}}{\mathrm{card}\mathcal{R}(\mathbf{x})} \tag{42}$$

  which is the resulting distribution of the following procedure :

  1. choose a pair $\{v, w\}$ of $\mathcal{R}(\mathbf{x})$,
  2. choose v or w with probability 0.5.

4.3.2 Toy example

Let present a toy example to see how this kind of kernel can be implemented.

We consider the natural parametrization describing rectangles :

$$S = K \times M \subset \mathbb{R}^2 \times \mathbb{R}^3 \qquad u = \begin{pmatrix} x \\ y \\ \theta \\ L \\ l \end{pmatrix}$$

Let consider the following symmetrical relation :

$$u \sim v \iff \|u-v\|_K \leq d_{\max} \quad \text{with} \quad \|u\|_K = \sqrt{x^2+y^2}$$

To generate z, we thus need to include an underlying ball in Σ.

First parametrization

When dealing with balls an intuitive parametrization is the circular one :

$$\Sigma = [0, d_{\max}] \times [0, 2\pi] \times M \qquad z = \begin{pmatrix} \rho \\ \varphi \\ \theta \\ L \\ l \end{pmatrix} \qquad \eta_u(z) = \begin{pmatrix} x_u + \rho\cos(\varphi) \\ y_u + \rho\sin(\varphi) \\ \theta \\ L \\ l \end{pmatrix}$$

The problem of such a parametrization is that the Jacobian is given by :

$$|J_{\eta_u^{-1}}(v)| = \frac{1}{\rho}$$

which leads to the following birth ratio :

$$R(\mathbf{x}, \mathbf{x} \cup v) = \frac{h(\mathbf{x} \cup v)}{h(\mathbf{x})} \frac{p_d}{p_b} \frac{1}{2\pi * d_{\max}} \frac{j_d^{\mathbf{x} \cup v}(v)}{\sum_{u \in \mathbf{x}} j_b^{\mathbf{x}}(u) \, \frac{\mathbf{1}(\|u-v\|_2 \leq d_{\max})}{\|u-v\|_2}}$$

We thus have to face the classical problem of diffderentiability of the circular parametrization. It is problematic, because as we will see later (cf. section 5) we need some bounds on Green's ratio to obtain Harris recurrence and geometric ergodicity.

Second parametrization

Let consider :

$$\Sigma = \{x \in \mathbb{R}^2 \ \ \|x\|_2 \leq d_{\max}\} \times M \qquad z = \begin{pmatrix} x \\ y \\ \theta \\ L \\ l \end{pmatrix} \qquad \eta_u(z) = \begin{pmatrix} x_u + x \\ y_u + y \\ \theta \\ L \\ l \end{pmatrix}$$

The Jacobian is obviously equal to 1, and using the uniform distribution on Σ, the birth ratio can be written :

$$R(\mathbf{x}, \mathbf{x} \cup v) = \frac{h(\mathbf{x} \cup v)}{h(\mathbf{x})} \frac{p_d}{p_b} \frac{1}{\pi * d_{\max}^2} \frac{j_d^{\mathbf{x} \cup v}(v)}{\sum_{u \in \mathbf{x}} j_b^{\mathbf{x}}(u) \, \mathbf{1}(\|u-v\|_2 \leq d_{\max})}$$

The problem shifts to the uniform generation in a disk. To achieve it, a good solution is to generate points uniformly in the square of width $2d_{\max}$ until one of them falls in the disk.

5 Improved MCMC point process sampler

We recall we want to sample from a point process distribution $\pi(.)$ defined by :

- a Poisson point process distribution $\mu(.)$ defined by an intensity measure $\nu(.)$,
- a density $h(.)$ defined with respect to this Poisson point process.

We extend Geyer and Møller algorithm for point process with an inhomogeneous intensity measure $\nu(.)$ and focus on an algorithm that uses a proposition kernel which is a linear combination of the previously described proposition kernels. We recall that we write the proposition kernel as :

$$Q(\mathbf{x}, A) = \sum_m Q_m(\mathbf{x}, A) \text{ with } Q(\mathbf{x}, \mathcal{C}) \leq 1 \tag{43}$$

and we note

$$p_m(\mathbf{x}) = Q_m(\mathbf{x}, \mathcal{C}) \tag{44}$$

the probability of choosing Q_m while in state $\mathbf{x}$.

5.1 Algorithm

We distinguish three kinds of kernels.

5.1.1 Birth or death

We note birth or death kernel Q_{BD}. We recall that the two associated Green's ratio are :

$$R_{BD}(\mathbf{x}, \mathbf{x} \cup u) = \frac{p_d(\mathbf{x} \cup u)}{p_b(\mathbf{x})} \frac{h(\mathbf{x} \cup u)}{h(\mathbf{x})} \frac{\nu(S)}{n(\mathbf{x}) + 1} \tag{45}$$

and

$$R_{BD}(\mathbf{x}, \mathbf{x} \setminus u) = \frac{p_b(\mathbf{x} \setminus u)}{p_d(\mathbf{x})} \frac{h(\mathbf{x} \setminus u)}{h(\mathbf{x})} \frac{n(\mathbf{x})}{\nu(S)} \tag{46}$$

This transformation do not add any condition on $\nu(.)$

5.1.2 Non jumping transformations

Without any loss of generality, we can suppose that there is only one non jumping kernel. We denote it by Q_{NJ}. The associated ratio is given by :

$$y = \mathbf{x} \setminus u \cup v \qquad R_{NJ}(\mathbf{x}, \mathbf{y}) = \frac{h(\mathbf{y})}{h(\mathbf{x})} \frac{j^{\mathbf{y}}(v)}{j^{\mathbf{x}}(u)} \frac{f_Z^{(\mathbf{y},v)}(\tilde{z})}{f_Z^{(\mathbf{x},u)}(z)} \tag{47}$$

where $j^{\mathbf{x}}(u)$ is the probability of choosing point u in $\mathbf{x}$, and $f(.)$ is the density of the associated auxiliary random variable Z that is used to sample the new object v.

5.1.3 Birth or death in a neighborhood

We denote this kind of kernel Q_{BDN}. We have seen that a general expression of associated Green's ratio is given by :

$$\mathbf{y} = \mathbf{x} \cup v \qquad v = \eta_u(z) \qquad v \in V(u) \tag{48}$$

In case of a birth :

$$R_{BDN}(\mathbf{x}, \mathbf{x} \cup v) = \frac{h(\mathbf{x} \cup v)}{h(\mathbf{x})} \frac{p_d}{p_b} \frac{j_d^{\mathbf{x} \cup v}(v)\ f_\nu(v)}{\sum_{u \in \mathbf{x}} j_b^{\mathbf{x}}(u)\ f_Z(\eta_u^{-1}(v))\Lambda_u(v)} \tag{49}$$

while in case of a death :

$$R_{BDN}(\mathbf{x}, \mathbf{x} \setminus v) = \frac{h(\mathbf{x} \setminus v)}{h(\mathbf{x})} \frac{p_b}{p_d} \frac{\sum_{u \in \mathbf{x} \setminus v} j_b^{\mathbf{x} \setminus v}(u)\ f_Z(\eta_u^{-1}(v))\ \Lambda_u(v)}{j_d^{\mathbf{x}}(v)\ f_\nu(v)} \tag{50}$$

where :

- $v = \eta_u(z)$, η_u being a diffeomorphism that allows, using a random variable Z and one of its realization $z \in \Sigma$ to generate a new point in a neighborhood of u,
- $f_Z(.)$ is the density of the random variable with respect to Lebesgue measure on Σ,
- $\Lambda_u(.)$ is a prolongation of the Jacobian of η_u^{-1} on Σ,
- and $f_\nu(.)$ is the density of intensity measure $\nu(.)$ with respect to Lebesgue measure on S.

We recall that, for birth or death in a neighborhood, we have taken the following discrete distributions on objects :

for a birth :

$$j_b^{\mathbf{x}}(u) = \frac{1}{n(\mathbf{x})} \tag{51}$$

which, from a practical point of view, means :

1. choose **uniformly** one object u in the current configuration
2. propose to add an object $v \in V(u)$.

for a death :

$$j_d^{\mathbf{x}}(u) = \frac{\frac{1}{2} * \operatorname{card}\{\{v,w\} \in \mathcal{R}(\mathbf{x}) \qquad u \in \{v,w\}\}}{\operatorname{card}\mathcal{R}(\mathbf{x})} \tag{52}$$

which is the resulting distribution of the following procedure :

1. choose a pair $\{v, w\}$ of $\mathcal{R}(\mathbf{x})$,
2. choose v or w with probability 0.5.
3. propose to remove the chosen object.

5.2 Algorithm

Let detail the algorithm we propose :

Algorithm B

For a given state $X_t = \mathbf{x}$

1. *choose one of the previously described proposition kernels $Q_m(.,.)$ with probability $p_m(\mathbf{x})$,*
2. *sample $\mathbf{y}$ according to the chosen kernel : $\mathbf{y} \sim Q_m(\mathbf{x}, .)$,*
3. *compute associated Green's ratio $R_m(\mathbf{x}, \mathbf{y})$ and acceptance rate $\alpha_m(\mathbf{x}, \mathbf{y})$,*
4. *accept the proposition $X_{t+1} = \mathbf{y}$ with probability α, and reject it otherwise*

5.3 Conditions

5.3.1 Usual stability condition

We first recall the usual stability condition :

Condition 1

A point process with unnormalized density $h(.)$ with respect to $\mu(.)$ is stable if there exists a real number R_h such that :

$$h(\mathbf{x} \cup u) \leq R_h h(\mathbf{x}) \ \ \forall \mathbf{x} \in \mathcal{C}, \ \ \forall u \in S$$

5.3.2 Mixture conditions

We then add three conditions on the mixture of proposition kernels :

Condition 2 *We suppose that*

1. *the probability of choosing a transformation $p_m(\mathbf{x})$ do not depend on $\mathbf{x}$ (ie. is constant),*
2. *in "birth or death" and "birth or death in a neighborhood" updates the probabilities of choosing a birth p_b and a death p_d do not depend on $\mathbf{x}$*
3. *the probability of proposing to do nothing is strictly positive :*

$$\forall \mathbf{x} \quad \mathbb{P}(\mathbf{y} = \mathbf{x}) \geq p_I > 0 \tag{53}$$

Some comments have to be made : the two first conditions are useful because they simplify the proofs of convergence. It is however possible to remove them and to keep good properties of convergence of the algorithm.

The last condition obviously implies the aperiodicity of the Markov chain.

5.3.3 Bounds on birth or death in a neighborhood

Next conditions are easy to ensure, and give useful bounds on the Green ratio of Birth or Death in a Neighborhood :

Condition 3 *We assume that the following inequalities hold :*

intensity

$$\exists (r_\nu, R_\nu) \in \mathbb{R}^2 \quad s.t. \quad 0 < r_\nu < f_\nu(.) < R_\nu \tag{54}$$

auxiliary random variable

$$\exists (r_Z, R_Z) \in \mathbb{R}^2 \quad s.t. \quad 0 < r_Z < f_Z(.) < R_Z \tag{55}$$

Jacobian

$$\exists (r_J, R_J) \in \mathbb{R}^2 \quad s.t. \quad 0 < r_J < |J_{\eta_u}^{-1}(.)| < R_J \tag{56}$$

The first condition is easy to obtain, since we are working on a bounded $S \subseteq \mathbb{R}^d$. The lower bound condition can be removed.

The second condition is related to the random variable that is used to generate a new object in the neighborhood of another one. Using a uniform distribution is useful in practice (it is easy to sample from) and gives the above condition.

The last condition is related to the differentiability of the transformation η_u. We already have seen (see section 4.3.2) on a toy example how to obtain this condition.

These two last conditions can be replaced by a single condition of minoration of the product $f_Z(z)|J^{-1}_{\eta_u^{-1}}(z)|$ (upper bounds are not needed).

5.3.4 Conditions on the relation-ship

We finally add the following condition on the relation-ship $\sim$ between objects :

Condition 4 *It is possible to divide S into a finite partition*

$$\exists k \in \mathbb{N} \quad s.t. \quad S = S_1 \cup \cdots \cup S_k \quad S_i \cap S_j = \emptyset \tag{57}$$

where every S_i verify the following property :

$$\forall i \in \{1, \dots, \mathbb{N}\}, \qquad \textit{if} \left\{ \begin{array}{l} u \in S_i \\ v \in S_i \end{array} \right. \quad \textit{then} \ u \sim v \tag{58}$$

Since S is bounded, this condition can be replaced by the following one :

Condition 4b **Non negligeable neighborhood**

$$\exists r_\sim \quad s.t. \quad \forall u \in S \qquad \lambda_S\left(V(u)\right) > r_\sim \tag{59}$$

If this condition holds, condition (4) also holds. It is sufficient to remark that since S is bounded, by adding points it is possible to iteratively build a sequence of strictly decreasing sets .

5.4 Property

The Markov Chain we built exhibits the required properties :

Proposition | *Algorithm B builds a Markov Chain that is $\pi(.)$ reversible and under Conditions 1,2,3 and 4, the Markov Chain is aperiodic, π-irreducible, Harris recurrent and geometrically ergodic.*

5.5 Proof-$\pi(.)$ reversibility

The first part of the proof relies on the following proposition :

Proposition | *Algorithm B builds a Markov Chain that is $\pi(.)$ reversible*

Proof. Proposition kernels and associated acceptance ratio were designed to build a $\pi(.)$ reversible Markov Chain under Green's framework.

5.6 Proof-Stability properties

Derivation presented here are adaptations of Geyer's work (see [6]). Our goal is to show :

- that the kernels we designed keep stability properties of the Markov chain,
- that it is possible to use reference point processes that are not homogeneous ($\nu(.)$ is not necessary a Lebesgue measure).

5.6.1 φ-Irreducibility and small sets

φ-Irreducibility

A Markov chain $(X_n)_{n\geq 0}$ on χ is φ irreducible if φ is a non-zero measure on χ and if for all $x \in \chi$ and $B \subset \chi$ such that $\varphi(B) > 0$, there exists an integer n, such that the probability $\mathbf{P}^n(x, B)$ of hitting B while starting in x is strictly positive : $\mathbf{P}^n(x, B) > 0$.

Here, $\mathbf{P}^n(x, B) = \mathbf{P}(X_{n+1} \in B | X_1 = x)$.

Small sets

A set C is small if there exists a non zero measure φ' and an integer n such that :

$$\mathbf{P}^n(x,B) \geq \varphi'(B) \qquad \forall x \in C \qquad \text{and } \; B \in \mathcal{B} \tag{60}$$

First result

We extend Geyer's result :

Proposition | *If the unnormalized density satisfies Conditions 1 and 2, then algorithm B simulates a φ-irreducible Markov chain, and every bounded set is small.*

Proof

We first define $\varphi^0(.)$ being the following nonzero measure on $\mathcal{C}$:

$$\forall \mathbf{x} \in \mathcal{C} \quad \begin{cases} \varphi^0(\mathbf{x}) = 1 & \text{if } \mathbf{x} = \emptyset \\ \varphi^0(\mathbf{x}) = 0 & \text{else} \end{cases} \tag{61}$$

It is possible to choose R_h satisfying **condition 1** and large enough so that :

$$\begin{array}{rcl} R_{BD}(\mathbf{x}, \mathbf{x} \setminus v) & \geq & \dfrac{1}{R_h\ \nu(S)} \dfrac{p_b}{p_d} \\ & \text{and} & \\ 1 & \geq & \dfrac{1}{R\ \nu(S)} \dfrac{p_b}{p_d} \end{array} \tag{62}$$

We now consider a given configuration $\mathbf{x}$, and a integer $m \geq n(x)$. Thus we have :

$$\mathbb{P}^m(\mathbf{x}, \{\emptyset\}) \geq \mathbb{P}^{n(\mathbf{x})}(\mathbf{x}, \{\emptyset\})\, \mathbb{P}^{m-n(\mathbf{x})}(\emptyset, \{\emptyset\}) \tag{63a}$$

$$\geq p_{BD}^{n(\mathbf{x})}\ p_d^{n(\mathbf{x})} \left(\frac{1}{R_h\ \nu(S)} \frac{p_b}{p_d} \right)^{n(\mathbf{x})} p_I^{m-n(\mathbf{x})} \tag{63b}$$

$$\geq p_{BD}^{n(\mathbf{x})}\ p_b^{n(\mathbf{x})}\ p_I^{m-n(\mathbf{x})} \left(\frac{1}{R_h\ \nu(S)} \right)^{n(\mathbf{x})} \tag{63c}$$

Which finally leads to :

$$\mathbb{P}^m(\mathbf{x}, \{\emptyset\}) \geq p_{BD}^m\ p_b^m\ p_I^m \left(\frac{1}{R_h\ \nu(S)} \right)^m \tag{64}$$

This shows that the Markov chain is φ^0-irreducible since $\mathbb{P}^m(\mathbf{x}, \{\emptyset\}) > 0$ when $m \geq n(x)$. The same calculation shows that every bounded set is small. For a given m, let consider :

$$C = \{\mathbf{x} \in \mathcal{C} : n(\mathbf{x}) \leq m\} \tag{65}$$

Introduce

$$c = p_{BD}^m \, p_b^m \, p_I^m \, \left(\frac{1}{R_h \, \nu(S)} \right)^m \tag{66}$$

and thus equation (60) is established with $\varphi' = c\varphi^0$.

As pointed out by Geyer ([6]), in Meyn and Tweedie's work there are two different notions : small set and *petite set*. Since our Markov chain is aperiodic (condition 3.3), these two notions are equivalent.

5.6.2 Harris recurrence and geometric ergodicity

Harris Recurrence

Dealing with a Markov chain that has a stationary distribution $\pi()$, Harris recurrence is the following property :

For all $\mathbf{x} \in \mathcal{C}$, and all π positive set A, there exists n such that : $P^n(\mathbf{x}, A) = 1$ for any $\mathbf{x} \in \mathcal{C}$.

A practically convenient way of showing Harris Recurrence is to use a drift condition. For a given function W, define :

$$\mathbf{P}W(\mathbf{x}) = \mathbb{E}[W(X_{t+1})|X_t = \mathbf{x}] = \int \mathbf{P}(\mathbf{x}, d\mathbf{y})W(\mathbf{y}) \tag{67}$$

W is said to be **unbounded** off small sets if for every $\gamma > 0$, the level set $\{\mathbf{x} \in \mathcal{C} : W(\mathbf{x}) \leq \gamma\}$ is small.

A Markov Chain satisfies the **drift condition for recurrence** if there exists :

- a function $W : \mathcal{C} \rightarrow (0, \infty)$ which is unbounded off small sets,
- and a small set C such that

$$\mathbf{P}W(\mathbf{x}) \leq W(\mathbf{x}) \qquad \mathbf{x} \notin C \tag{68}$$

If a chain satisfies the drift condition for recurrence, then it is Harris recurrent (see [12]).

Geometric Ergodicity

A Markov Chain is **geometrically ergodic** if there exists a constant $r > 1$ such that :

$$\sum_{n=1}^{\infty} r^n \|\mathbb{P}^n(\mathbf{x}, .) - \pi(.)\|_{TV} < \infty \qquad \forall \mathbf{x} \in \mathcal{C} \tag{69}$$

This is implied (see [12]) by the **geometric drift condition** :

There exists a function $W : \mathcal{C} \to [1, \infty)$, constants $b < \infty$ and $\varpi < 1$, and a small set C such that :

$$\mathbf{P}W(\mathbf{x}) \leq \varpi W(\mathbf{x}) + b\mathbf{1}_C(\mathbf{x}) \qquad \forall \mathbf{x} \in \mathcal{C} \tag{70}$$

If the chain is aperiodic, Geyer pointed out that any W satisfying the geometric drift condition is unbounded off small sets. In other terms, geometric drift condition implies the drift condition for recurrence.

Proposition *Under conditions 1,2,3 and 4, Algorithm B simulates a Markov chain that is*

- *Harris recurrent*
- *and geometrically ergodic.*

Proof : We start the proof using the same function as Geyer : $W(\mathbf{x}) = A^{n(x)}$, with an arbitrary $A > 1$. There are two transformations that can add a point to the current configuration $\mathbf{x}$, birth or death and birth or death in a neighborhood.

Let denote by $\alpha^+_{BD}(\mathbf{x}, \mathbf{y})$ the probability of accepting a birth using the usual birth and death. Using conditions 1 and 2, and the expression of Green's ratio we obtain that :

$$\alpha^+_{BD}(\mathbf{x}, \mathbf{y}) \leq \frac{p_d}{p_b} \frac{R_h \, \nu(S)}{n(\mathbf{x}) + 1} \tag{71}$$

Thus for a given $\epsilon \in (0, 1)$, there exists K^{BD}_ϵ depending on R_h, p_b,p_d and $\nu(S)$ such that

$$\alpha^+_{BD}(\mathbf{x}, \mathbf{y}) < \epsilon \text{ when } n(\mathbf{x}) \geq K^{BD}_\epsilon \tag{72}$$

The expression of the ratio for a death also give that the probability of accepting a death is equal to one if $n(\mathbf{x})$ is large enough :

$$\alpha^-_{BD}(\mathbf{x}, \mathbf{y}) = 1 \text{ when } n(\mathbf{x}) \geq K^{BD}_\epsilon \tag{73}$$

We now denote by $\alpha^+_{BDN}(\mathbf{x}, \mathbf{y})$ the probability of accepting a birth using birth or death in a neighborhood. Using conditions 1,2 and 3 :

$$\alpha^+_{BDN}(\mathbf{x}, \mathbf{x} \cup v) \leq \frac{p_d}{p_b} \frac{R_h \, R_\nu}{r_Z r_J} \frac{j_d^{\mathbf{x} \cup v}(v)}{\sum_{u \in \mathbf{x}} j_b^{\mathbf{x}}(u) \mathbf{1}(u \in V(v))} \tag{74}$$

The term $\mathbf{1}(u \in V(v))$ comes from the expression of $\Lambda_u(v)$. We now focus on the ratio $j_d^{\mathbf{x}\cup v}(v)/\sum_{u\in\mathbf{x}} j_b^{\mathbf{x}}(u)\mathbf{1}(u \in V(v))$ in order to show it tends to zero as $n(\mathbf{x})$ tends to infinity. To achieve this we first use the expressions of j_d and j_b given by (52) and (51). We obtain :

$$\frac{j_d^{\mathbf{x}\cup v}(v)}{\sum_{u\in\mathbf{x}} j_b^{\mathbf{x}}(u)\mathbf{1}(u \in V(v))} = \frac{\frac{1}{2} * \frac{\text{card}\left\{\{u_i,u_j\} \in \mathcal{R}(\mathbf{x}\cup v) \text{ s.t. } v \in \{u_i,u_j\}\right\}}{\text{card}\mathcal{R}(\mathbf{x}\cup v)}}{\frac{\text{card}\left\{\{u_i,u_j\} \in \mathcal{R}(\mathbf{x}\cup v) \text{ s.t. } v \in \{u_i,u_j\}\right\}}{n(\mathbf{x})}} \tag{75}$$

If we note $s(\mathbf{x})$ the number of interacting pair of points in $\mathbf{x}$, ie $s(\mathbf{x}) = \text{card}\{\mathcal{R}(\mathbf{x})\}$, we obtain, after calculus :

$$\boxed{\frac{j_d^{\mathbf{x}\cup v}(v)}{\sum_{u\in\mathbf{x}} j_b^{\mathbf{x}}(u)\mathbf{1}(u \in V(v))} = \frac{1}{2}\frac{n(\mathbf{x})}{s(\mathbf{x}+1)}} \tag{76}$$

This result is really important, since as detailed below, the ratio of number of points divided by the number of interacting points is going to zero, because of the quadratic behavior of $s(\mathbf{x})$.

Lemma

Under condition 4,

$$\sup\left\{\frac{n(\mathbf{x})}{s(\mathbf{x})} \quad : \mathbf{x} \;\; s.t. \;\; n(\mathbf{x}) = n\right\} \tag{77}$$

tends to zero as n tends to infinity.

Proof of the lemma :

Using condition 4, we partition S into k subsets S_i such that if two points belong to the same subset there are in relationship.

A given configuration $\mathbf{x}$ put n_i points in S_i so :

$$n(\mathbf{x}) = \sum_{i=1}^{k} n_i \qquad s(\mathbf{x}) \geq \sum_{i=1}^{k} \binom{n_i}{2} = s(n_1,\dots,n_k) \tag{78}$$

Considering the k-tuples $(n_1,\dots,n_k)$ in $\mathbb{R}^k$, it is well known that the minimum of $s(n_1,\dots,n_k)$ under constraint $\sum n_i = n$ is achieved when $n_i = n/k$ for every i. Thus we obtain that :

$$s(n_1,\dots,n_k) \geq k * \frac{(n/k)*(n/k-1)}{2} \tag{79}$$

which leads to :

$$\text{if } n(\mathbf{x}) \geq k \qquad s(\mathbf{x}) \geq \frac{n(\mathbf{x})(n(\mathbf{x})-k)}{2k} \tag{80}$$

because of this quadratic behavior of s, the lemma is shown.

We now can conclude on the behavior of α^+_{BDN} :

For a given $\epsilon \in (0,1)$, there exists K^{BDN}_ϵ depending on R_h, k, R_ν, r_Z, r_J, p_b, p_d and $\nu(S)$ such that

$$\alpha^+_{BDN}(\mathbf{x}, \mathbf{x} \cup v) < \epsilon \ \text{ when } \ n(\mathbf{x}) \geq K^{BDN}_\epsilon \tag{81}$$

and by symmetry of the ratio if $n(\mathbf{x})$ is large enough, the probability of accepting a death in a neighborhood is equal to 1 :

$$\alpha^-_{BDN} = 1 \ \text{ when } \ n(\mathbf{x}) \geq K^{BDN}_\epsilon \tag{82}$$

Let note p_{st} the probability of doing a move that let the number of points unchanged. p_{st} involves :

- probability of proposing a birth (as usual or in a neighborhood) and rejecting the proposition,
- probability of proposing a death (as usual or in a neighborhood) and rejecting the proposition,
- probability of proposing a non-jumping transformation.

Thus if we define $K_\epsilon = \max(K^{BD}_\epsilon, K^{BDN}_\epsilon)$, we obtain that for $n(\mathbf{x}) \geq K_\epsilon$

$$\begin{aligned}
\mathbf{P}W(\mathbf{x}) &= \mathbb{E}[W(X_{t+1})|X_t = \mathbf{x}] \\
&\leq A^{n(\mathbf{x})+1}(p_{BD}\ p_b\ \alpha^+_{BD} + p_{BDN}\ p_b\ \alpha^+_{BDN}) + \ldots \\
&\qquad \ldots A^{n(\mathbf{x})} p_{st} + A^{n(\mathbf{x})-1}(p_{BD}\ p_d \alpha^-_{BD} + p_{BDN}\ p_d \alpha^-_{BDN}) \\
&\leq \left(A\epsilon(p_{BD}\ p_b + p_{BDN}\ p_b)) + p_{st} + \frac{1}{A}(p_{BDN}\ p_d + p_{BD}\ p_d) \right) W(\mathbf{x})
\end{aligned}$$

Let fix ϵ small enough to obtain a $\varpi < 1$ such that :

$$\mathbf{P}W(\mathbf{x}) \leq \varpi W(\mathbf{x}) \qquad \text{for} \quad n(\mathbf{x}) \geq K_\epsilon \tag{83}$$

by taking :

$$\varpi = A\epsilon(p_{BD}\ p_b + p_{BDN}\ p_b)) + p_{st} + \frac{1}{A}(p_{BDN}\ p_d + p_{BD}\ p_d) \tag{84}$$

We now consider $C = \{\mathbf{x} \in \mathcal{C} : n(\mathbf{x}) < K_\epsilon\}$ which is small. It is obvious to see that :

$$\mathbf{P}W(\mathbf{x}) \leq A^{K_\epsilon+1} \qquad \text{for} \quad \mathbf{x} \in C \tag{85}$$

and thus, taking $b = A^{K_\epsilon+1}$ the geometric drift condition (70) is achieved.

6 Experiments

6.1 Poisson point process

We consider configurations of rectangles :

$$M = [-\frac{\pi}{2}, \frac{\pi}{2}] \times [L_{\min}, L_{\max}] \times [l_{\min}, l_{\max}]$$

and take the unit square of $\mathbb{R}^2$ for K :

$$K = [0,1] \times [0,1]$$

We use for the reference intensity measure :

$$\nu(.) = (\lambda_K \times \frac{\lambda_M}{\lambda_M(M)})(.) \tag{86}$$

In the following sections, we consider a Poisson point process X with density $h(.)$ against the reference Poisson point process of intensity $\nu(.)$ where :

$$h(\mathbf{x}) \propto \exp(\sum_{u \in \mathbf{x}} \beta(u)) \tag{87}$$

According to well known results (see [2] for instance), for a given Borel set $A \subset S$, the random variable $N_A(X)$ which counts the number of points of X falling in A is Poisson distributed as follow :

$$N_A(X) \qquad \text{Poisson distributed with mean} \qquad \mathbb{E}[N_A(X)] = \int_A \beta(u) d\nu(u) \tag{88}$$

6.2 Target distribution and Total Variation

We present here a way to check that the algorithm converges to the desired distribution.

We propose to choose one or more A and to verify property (88) over a number N of trajectories by computing an upper bound of the total variation distance between the empirical law of $N(X_t)$ at time t and the target distribution.

Let $(\hat{p}^t_{(n,A)})_{n \geq 0}$ be, the empirical discrete distribution of the number of points of (X_t) falling in A :

$$\hat{p}^t_n = \frac{1}{N} \sum_{i=1}^{N} \mathbf{1}(N_A(X^i_t) = n) \tag{89}$$

If (p_n) is the theoretical target distribution (given by equation (88)), we consider :

$$d^{m_{max}}_{m_{min}}(t, A) = \sum_{n=m_{min}}^{m_{max}} |p_n - \hat{p}^t_n| \tag{90}$$

where m_{min}, m_{max} are two relevant parameters depending on the target Poisson distribution that define a truncated support on which empirical and theoretical distributions are compared.

6.3 Birth or death

We first test the convergence of the Markov Chain with only the birth or death kernel.

In the first experiment, we use $A = S$. We take $\beta(u) = e$ so that the number of points $n(X)$ should follow a Poisson distribution with mean e since $\nu(S) = 1$.

Figure 3 shows a result for $e = 30$. We have run $N = 10000$ simulations and taken $m_{min} = 15$ and $m_{max} = 45$. The first plot shows d as a function of time t. We verify on this plot the good behavior of the discrete distributions.

Figure 4 shows three results for $e = 30$, $e = 15$ and $e = 5$, the latter being interesting because it shows that the border $n = 0$ has no effect on the simulation quality.

The support we took to measure the distance d between the empirical and the theoretical distributions were :

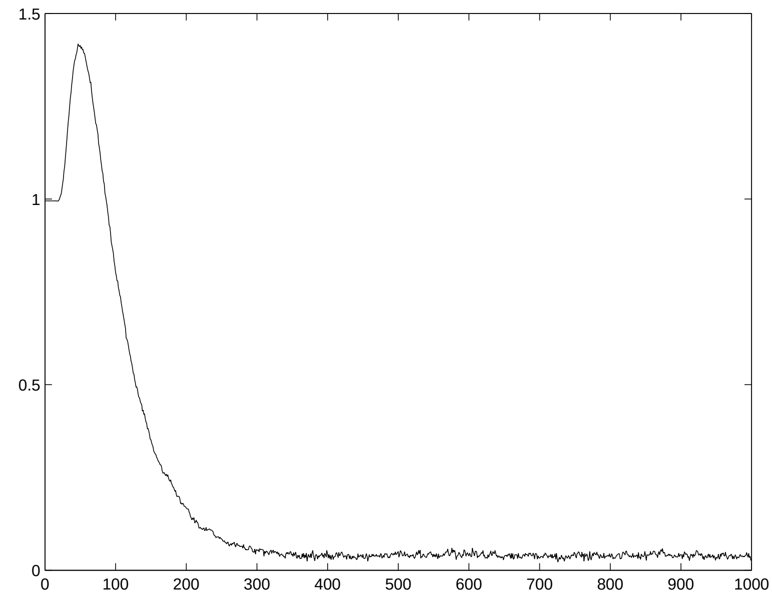
e	m_{min}	m_{max}
5	0	30
15	0	30
30	15	45

These experiments show that the Markov chain $N_S(X_t)$ converges to the desired Poisson distribution on S. A point is missing here : we should verify, that given their number, points of a configuration are identically and uniformly distributed. However, this point is related to the pseudo-random generator used, and there exists a huge literature on how to check uniformity.

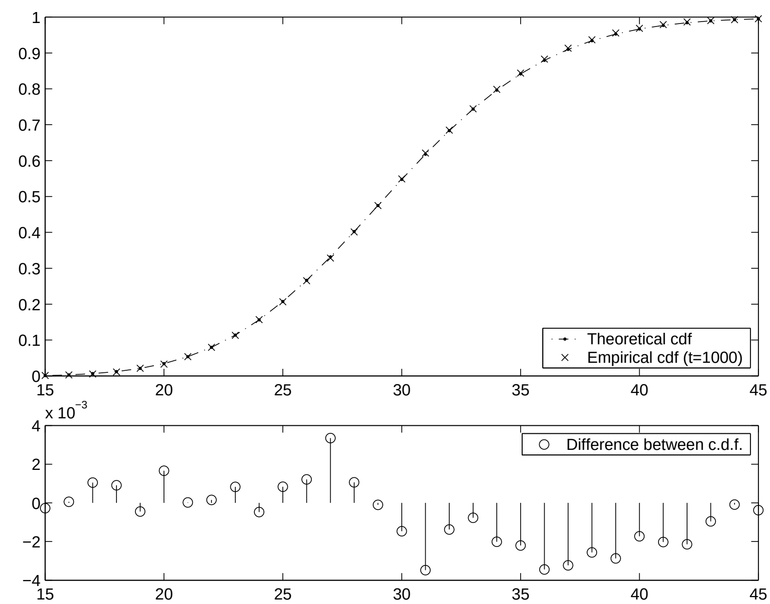
6.4 Birth and death in a neighborhood

We aim at showing in this section that the theoretical expression for birth or death in a neighborhood is valid and that this kind of transformation is useful.

To show that such an update is useful, we use a non homogeneous Poisson point process, and verify that the convergence rate is faster when using the update. Non homogeneous point process introduces a spatial correlation between points. Thus birth or death in a neighborhood should be useful to sample such point processes.



$d(t, S)$ (ordinate) depending on the number of iteration t (abscisa).



Comparison between theoretical cumulative distribution function (cdf) and empirical one of X_t when $t = 1000$ (abscissa : support of the distribution).

Figure 3: Result of first experiment, with $e = 30$ and $N = 10000$ trajectories.

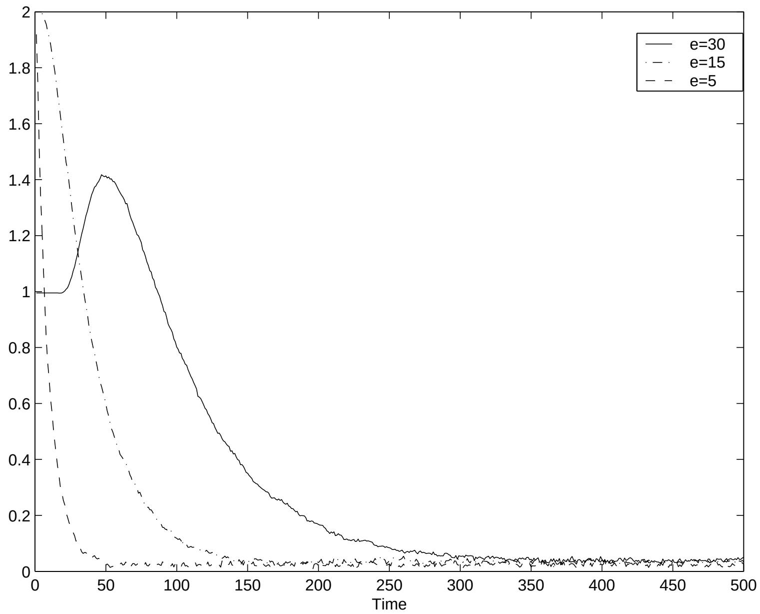


Figure 4: Distance d (ordinate) function of time for $e = 5$, $e = 15$ and $e = 30$ using $A = S$, $N = 10000$ runs for each Markov Chain. The starting point is always the empty configuration.

6.4.1 Non homogeneous Poisson point process

Let consider a Poisson point process on the compact set $[0,1] \times [0,1]$, and a non-constant intensity function :

$$\beta(u) = \begin{cases} \rho * e & \text{if } u \in S_{\text{sub}} = [0, \frac{1}{2}] \times [0, \frac{1}{2}] \\ e & \text{else} \end{cases} \tag{91}$$

This partition of S is described by figure 5. For $\rho = 1$, the Poisson point process is homo-

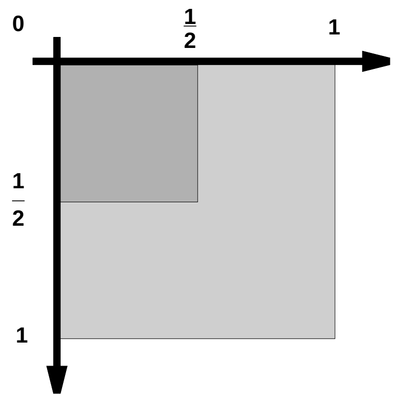


Figure 5: Partition of S used.

geneous, with intensity e. For $\rho > 1$, area S_{sub} is favored. The expected number of points is indeed given by :

$$\begin{aligned} \mathbb{E}[N_{S_{\text{sub}}}(X)] &= \rho \frac{e}{4} \\ \mathbb{E}[N_{S \setminus S_{\text{sub}}}(X)] &= \frac{3e}{4} \\ \mathbb{E}[N_S(X)] &= \frac{(\rho+3)e}{4} \end{aligned} \tag{92}$$

Figure 6 shows a realization of such an inhomogeneous point process using $e = 50$ and $\rho = 6$.

6.4.2 Experiments

We take $e = 20$, and look at several cases $(\rho = 1, \rho = 3, \rho = 6)$

For each case, we did two experiments, each of them consisting in 10000 trajectories :

- the first one using only usual birth or death proposition kernel,
- the second one using also birth or death in a close neighborhood ($p_{BDN} = 0.5$).

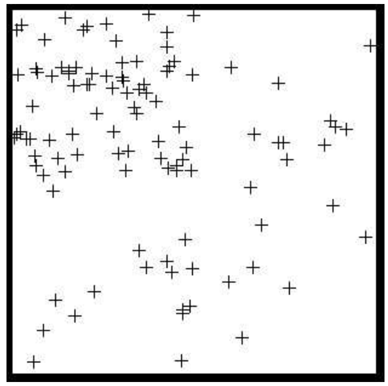


Figure 6: Realization of an inhomogeneous Poisson point process using $e = 50$ and $\rho = 6$.

6.4.3 Distributions of interest

As for the previously described experiments, we need some distance functions. We use two distance functions :

- $d_S = d(t, S_{\text{sub}})$, to verify that $N_{S_{\text{sub}}}$ ergodically follows a Poisson distribution with mean $\rho * 5$,
- $d_{S^c} = d(t, S \setminus S_{\text{sub}})$ to verify that the corresponding random variable is Poisson distributed with mean 15.

6.4.4 Birth or death in a neighborhood

We use a $\|.\|_\infty$ neighborhood :

$$u \sim v \iff \max(|x_u - x_v|, |y_u - y_v|) \leq d_{\max} \tag{93}$$

and a parameter : $d_{\max} = 0.1$. Figure 7 shows a realization of an inhomogeneous Poisson point process using $e = 20$ and $\rho = 6$. Points that are related are neighbors.

6.4.5 Results

Figures 8, 9 and 10 respectively show results with $\rho = 1$, $\rho = 3$ and $\rho = 6$.

These results show that the Markov Chain ergodically converges faster when BDN[1] is used,

[1] Birth or Death in a Neighborhood

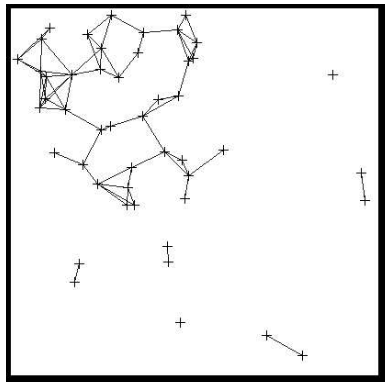


Figure 7: Realization of an inhomogeneous Poisson point process using $e = 20$ and $\rho = 6$. If the L^∞ distance between two points is less than 0.1, points are related on the plot.

if inhomogeneity is important enough. On figure 10, it is important to notice that BDN makes the convergence be faster on S_{sub} and slower on $S \setminus S_{sub}$.

This is a good behavior : it means exploration is improved in highly probable part of the space, while other parts are penalized.

Of course a complete study should take into account the mixing probabilities p_{BD}, p_{BDN}, p_b and the parameter d_{max}.

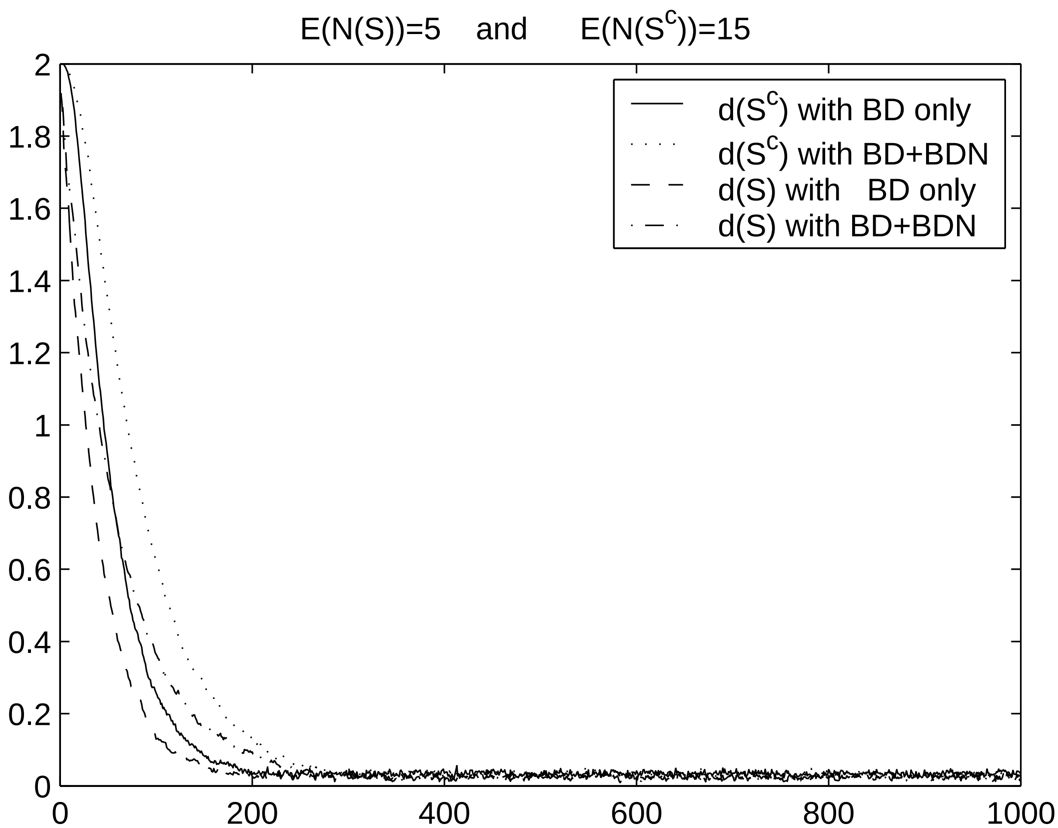


Figure 8: Distances d_S and d_{S^c} function of time with $\rho = 1$ using $N = 10000$ runs for two experiments : first one using only Birth or Death kernel (BD), second one using also Birth or Death in a Neighborhood (BDN).

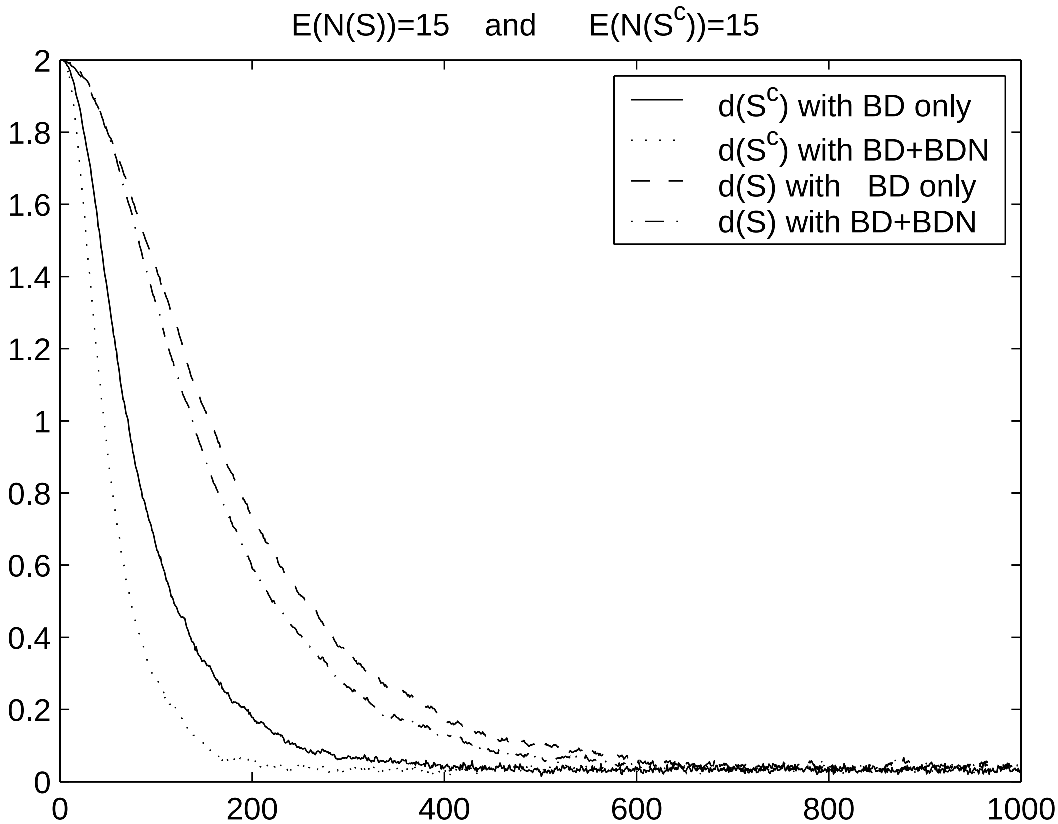


Figure 9: Distances d_S and d_{S^c} function of time with $\rho = 3$ using $N = 10000$ runs for two experiments : first one using only Birth or Death kernel , second one using also Birth or Death in a Neighborhood.

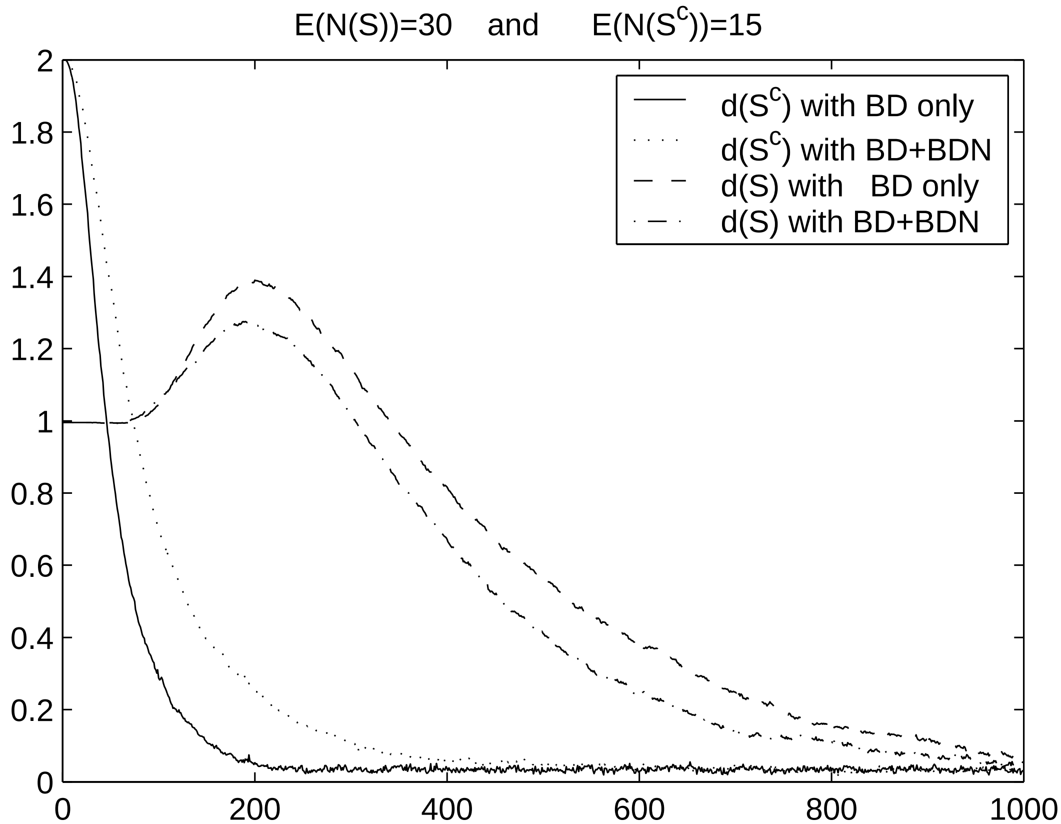


Figure 10: Distances d_S and d_{S^c} function of time with $\rho = 6$ using $N = 10000$ runs for two experiments : first one using only Birth or Death kernel, second one using also Birth or Death in a Neighborhood.

7 Improved birth and death in a neighborhood

7.1 Asymmetric relationships

We present here a kind of birth and death in a neighborhood that is useful for practical applications.

We consider an **intensity potential** $\gamma : S- > \mathbb{N}$. This potential quantifies some state of an object $u \in S$. For instance, we suppose that this potential can be equal to 0 or to 1.

In our building detection framework, this corresponds to the level of the data term. We suppose that $\gamma(u) = 1$ corresponds to a pertinent object with respect to the data. In practice, it is really convenient not to loose time exploring neighborhoods of an irrelevant object.

Thus we propose here a birth or death in a neighborhood kernel that only acts on relevant objects. It obeys to the following procedure :

- If birth has been selected, do :
  1. Choose an object in $\{u \in \mathbf{x} \,\text{s.t.}\, \gamma(u) = 1\} = \gamma_1(\mathbf{x})$,
  2. Propose an v in $V(u)$,
  3. Accept v with the corresponding ratio.
- If death has been selected :
  1. Choose a pair of $\{u, v\}$ objects in $\mathcal{R}(\mathbf{x})$ such that u or v belongs to $\gamma_1(\mathbf{x})$,
  2. if $\gamma(u) = 1$ and $\gamma(v) = 0$, propose $\mathbf{y} = \mathbf{x} \setminus v$, otherwise, with a probability 0.5 propose to remove u, and with probability 0.5, propose to remove v.

First, let remark that this kind of kernel can be described by birth or death in a neighborhood as described in section 4, using :

$$j_b^{\mathbf{x}}(u) = \frac{\mathbf{1}_{\gamma_1}(u)}{\text{card } \gamma_1(\mathbf{x})} \tag{94}$$

and

$$j_d^{\mathbf{x} \cup v}(v) = \frac{1}{\text{card } \mathcal{R}(\mathbf{x})} \sum_{u \,:\, \{u,v\} \in \mathcal{R}(\mathbf{x})} \frac{1}{2} \mathbf{1}_{\gamma_1}(u) \mathbf{1}_{\gamma_1}(v) + \mathbf{1}_{\gamma_1}(u)(1 - \mathbf{1}_{\gamma_1}(v)) \tag{95}$$

To keep stability properties, we should show that left handside of equation (76) still decreases to zero.

Using the same notations, we have :

$$\frac{j_d^{\mathbf{x}\cup v}(v)}{\sum_{u\in\mathbf{x}} j_b^{\mathbf{x}}(u)\mathbf{1}(u\in V(v))} \leq \frac{\mathrm{card}\gamma_1(\mathbf{x})}{s(\mathbf{x}\cup v)} \; \frac{\sum_{u\in x}\mathbf{1}_{\gamma_1}(u)[1-\frac{1}{2}\mathbf{1}_{\gamma_1}(v)]\mathbf{1}(u\in V(v))}{\sum_{u\in\mathbf{x}}\mathbf{1}_{\gamma_1}(u)\mathbf{1}(u\in V(v))}$$

which, using $\gamma_1(\mathbf{x}) \leq n(\mathbf{x})$ shows that the drift condition is still valid.

If we use the framework we used in the precedent section to show BDN, we propose to define γ as $\gamma(u) = \mathbf{1}(u \in S_{sub})$, thus BDN is only applied to a point living in S_{sub}. Figure 11 shows the result : this kind of transformation improves convergence rate on S_{sub} and penalize $S \setminus S_{sub}$.

7.2 Pre-explorative birth and death in a neighborhood

Presentation

Since death or birth of a point in a neighborhood allows to focus the proposition on some interesting part of the space to explore, it can be useful to use a pre-explorative scheme combined to birth in a neighborhood to improve the quality of the result and the exploration ability of the Markov chain.

Following the idea of the perturbation kernel presented in section 4, we propose :

1. to partition Σ into several Σ_i defined by their centers z_i,
2. to compute the energy variations $U(\mathbf{x} \cup \eta_u(z_i))$,
3. to select one of the z_i according to a discrete distribution p_i depending on these energy variations,
4. to generate z in the selected ball $B(z_i)$,
5. to propose $v = \eta_u(z)$,
6. to compute all the pairs of related points involving v : $\{v, w\}$, and for each of these pairs, to compute :
   - the associated Jacobian $\Lambda_w(v)$,
   - probability $p_I(w, v)$ of choosing the ball $B(z_i)$ such that $\eta_w^{-1}(v) \in B(z_i)$,
7. to put all together using the following ratio :

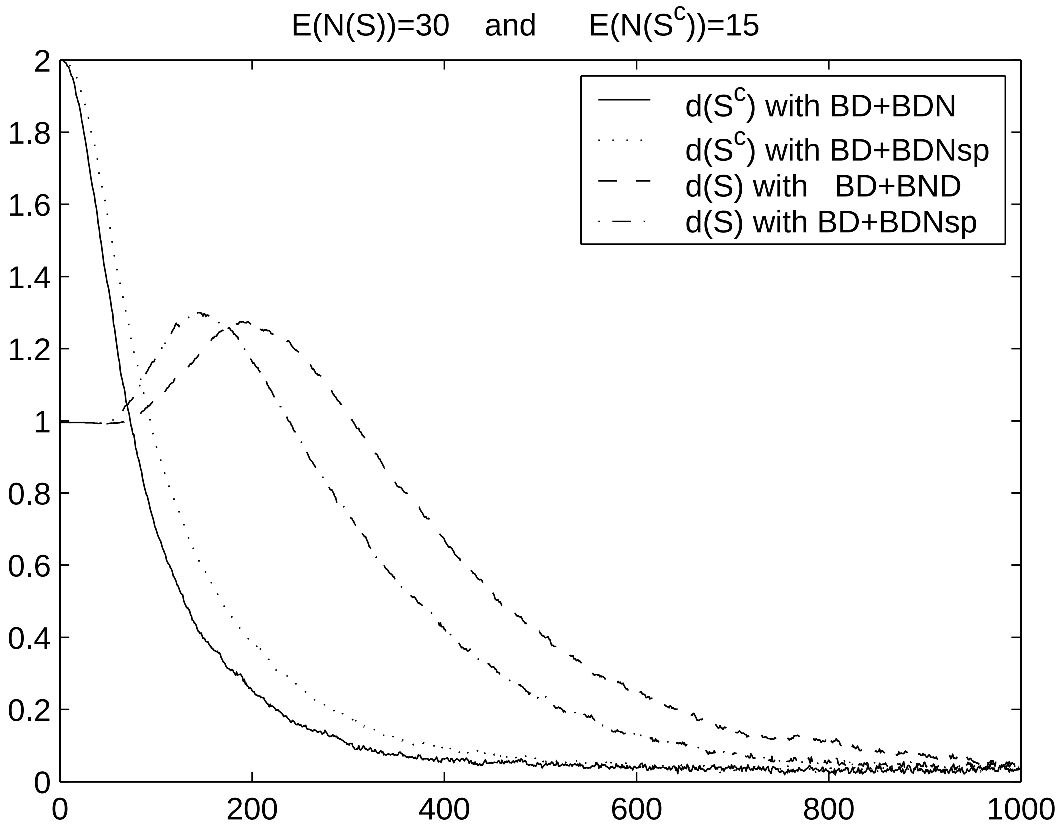


Figure 11: Distances d_S and d_{S^c} function of time with $\rho = 6$ using $N = 10000$ runs for two experiments : first one using only Birth or Death and Birth or Death in a a neighborhood, second one using also Birth or Death and Non symmetric Birth or Death in a Neighborhood.

$$\boxed{R(\mathbf{x}, \mathbf{x} \cup v) = \frac{h(\mathbf{x} \cup v)}{h(\mathbf{x})} \frac{p_d}{p_b} \frac{j_d^{\mathbf{x} \cup v}(v)\ f_\nu(v)}{\sum_{u \in \mathbf{x}} j_b^{\mathbf{x}}(u)\ \frac{p_I(u,v)}{\lambda(B(z_i))} \Lambda_u(v)}} \tag{96}$$

while when deleting a point :

1. choose at random a pair $\{u, v\}$ of $\mathcal{R}(\mathbf{x})$, and choose one of the object of the pair with probability $\frac{1}{2}$,
2. once u is chosen, to compute all pairs $\{u, v\}$ involving u, and for each of them, to compute :
   - the associated Jacobian $\Lambda_v(u)$,
   - probability $p_I(v, u)$ of choosing the ball $B(z_i)$ such that $\eta_v^{-1}(u) \in B(z_i)$,
3. to put all together using the following ratio :

$$\boxed{R(\mathbf{x}, \mathbf{x} \setminus v) = \frac{h(\mathbf{x} \setminus v)}{h(\mathbf{x})} \frac{p_b}{p_d} \frac{\sum_{u \in \mathbf{x} \setminus v} j_b^{\mathbf{x} \setminus v}(u)\ \frac{p_I(u,v)}{\lambda(B(z_i))}\ \Lambda_u(v)}{j_d^{\mathbf{x}}(v)\ f_\nu(v)}} \tag{97}$$

The two ratio have been directly obtained by replacing $f_Z(z)$ by a mixture of uniform distributions on the ball $B(Z_i)$. It is worth noting that under condition 1 (usual stability) , we still have bounds needed by condition 3 and thus the Markov chain keeps its stability properties.

Conclusion

In this report we have presented some improvements to RJMCMC point process samplers adapted to image processing.

We have given a very general expressions of several updating schemes, like "pre-explorative" schemes and birth or death in a neighborhood. Experiments did show our theoretical results are valid.

We also presented some examples of useful improvements using our general expressions.

In our practical application (cf. [15]), this improvements proved to be efficient.

However, some additionnal work could be done on the kind of estimator used (MAP). An other direction that seems to be worth exploring is the influence of the reference measure on the quality of the result, since the exploration ability of the designed Markov chain is closely related to the quality of the simulated annealing procedure.

Acknowledgments

The authors would like to thank Marie Colette Van Lieshout and Radu Stoica from CWI for several interesting discussions. The work of the first author has been partially supported by the French Defense Agency (DGA) and CNRS.

References

[1] A. Baddeley and M. N. M. van Lieshout. Stochastic geometry models in high-level vision. In K.V. Mardia, editor, *Statistics and Images*, volume 1, pages 233–258. Abingdon: Carfax, 1993.

[2] O.E. Banorff-Nielsen, W.S Kendall, and M.N.M. van Lieshout, editors. *Stochastic Geometry Likehood and computation.* Chapman and Hall, 1999.

[3] S.P. Brooks and N. Friel. Classical model selection via simulated annealing. `http://www.statslab.cam.ac.uk/ mcmc/`, 2001.

[4] O. Cappé, C. P. Robert, and T. Rydén. Reversible jump MCMC converging to birth-and-death MCMC and more general continuous time sampler. `http://www.statslab.cam.ac.uk/ mcmc/`, 2002.

[5] L. Garcin, X. Descombes, J. Zerubia, and H. Le Men. Building detection by Markov object processes and a MCMC algorithm. *INRIA Research Report 4206*, June 2001.

[6] C. J. Geyer. *Stochastic Geometry Likehood and computation*, chapter Likehood Inference for Spatial Point Processes. Chapman and Hall, 1999.

[7] C.J. Geyer and J. Møller. Simulation and likehood inference for spatial point process. *Scandinavian Journal of Statistics*, Series B, 21:359–373, 1994.

[8] P.J. Green. Reversible jump Markov chain Monte-Carlo computation and Bayesian model determination. *Biometrika*, 57:97–109, 1995.

[9] M. Imberty and X. Descombes. Simulation de processus objets : Etude de faisabilité pour une application à la segmentation d'images. *INRIA Research Report 4516*, 2002.

[10] W. S. Kendall and J. Møller. Perfect Metropolis-Hastings simulation of locally stable spatial point processes. *Advances in Applied Probability (SGSA)*, 2000.

[11] C. Lacoste, X. Descombes, and J. Zerubia. A comparative study of point processes for line network extraction in remote sensing. *INRIA Research Report 4516*, 2002.

[12] S. P. Meyn and R.L. Tweedie. *Markov Chains and Stochastic Stability.* Springer-Verlag, London, 1993.

[13] M. Ortner. Extraction de caricatures de bâtiments sur des modèles numériques d'élévation. Master Thesis (DEA, in French), August 2001. `http://www-sop.inria.fr/ariana/personnel/Mathias.Ortner/rapport_DEA_ortner.ps.gz`.

[14] M. Ortner, X. Descombes, and J. Zerubia. Building detection from digital elevation models. *INRIA Research Report 4517*, July 2002.

[15] M. Ortner, X. Descombes, and J. Zerubia. Automatic 3d land register extraction from altimetric data in dense urban areas. *INRIA Research Report*, August 2003.

[16] M. Ortner, X. Descombes, and J. Zerubia. Building detection from digital elevation models. In *ICASSP*, volume III, Hong Kong, April 2003.

[17] H. Rue and M. Hurn. Bayesian object identification. *Biometrika*, 3:649–660, 1999.

[18] H. Rue and A. R. Syverseen. Bayesian object recognition with Baddeley's delta loss. *Adv. Appl. Prob*, 30:64–84, 1998.

[19] M. N. M. van Lieshout. *Markov Point Processes and their Applications.* Imperial College Press, London, 2000.

[20] M. N. M. van Lieshout and R. S. Stoica. Perfect simulation for marked point processes. *Research Report PNA-R0306, CWI, Amsterdam*, May 2003.